Supplement to "Panel experiments and dynamic causal effects: A finite population perspective"

(*Quantitative Economics*, Vol. 12, No. 4, November 2021, 1171–1196)

Iavor Bojinov
Technology and Operations Management Unit, Harvard Business School

Ashesh Rambachan
Department of Economics, Harvard University

Neil Shephard
Department of Economics and Department of Statistics, Harvard University

This supplement contains the following sections: Section A contains the proofs of the main results in the paper, Section B contains additional theoretical results discussed in the main text of the paper, Section C contains additional results from the simulation study, and Section D contains additional results for the empirical application to a panel experiment in experimental economics.

Appendix A: Proofs of main results

Proof of Theorem 3.1

We begin the proof with a lemma that will be used later on.

Lemma A.1. *Assume a potential outcome panel with an assignment mechanism that is individualistic* (*Definition* 3) *and probabilistic* (*Assumption* 2). *Define, for any* $\mathbf{w} \in \mathcal{W}^{(p+1)}$, *the random function* $Z_{i,t-p:t}(\mathbf{w}) := p_{i,t-p}(\mathbf{w})^{-1}\mathbb{1}\{W_{i,t-p:t} = \mathbf{w}\}$. *Then, over the assignment mechanism,* $\mathbb{E}(Z_{i,t-p:t}(\mathbf{w})|\mathcal{F}_{i,t-p-1}) = 1$ *and* $\mathrm{Var}(Z_{i,t-p:t}(\mathbf{w})|\mathcal{F}_{i,t-p-1}) = p_{i,t-p}(\mathbf{w})^{-1}(1 - p_{i,t-p}(\mathbf{w}))$, *and* $\mathrm{Cov}(Z_{i,t-p:t}(\mathbf{w}), Z_{i,t-p:t}(\tilde{\mathbf{w}})|\mathcal{F}_{i,t-p-1}) = -1$ *for all* $\mathbf{w} \neq \tilde{\mathbf{w}}$. *Moreover,* $Z_{i,t-p:t}(\mathbf{w})$ *and* $Z_{j,t-p:t}(\mathbf{w})$ *are, conditioning on* $\mathcal{F}_{1:N,t-p-1}$, *independent for* $i \neq j$.

Proof. The expectation is by construction, the variance comes from the variance of a Bernoulli trial. The conditional independence is by the individualistic assignment assumption. □

Iavor Bojinov: ibojinov@hbs.edu
Ashesh Rambachan: asheshr@g.harvard.edu
Neil Shephard: shephard@fas.harvard.edu

© 2021 The Authors. Licensed under the Creative Commons Attribution-NonCommercial License 4.0. Available at http://qeconomics.org. https://doi.org/10.3982/QE1744

For any $\mathbf{w}, \tilde{\mathbf{w}} \in \mathcal{W}^{(p+1)}$, let $u_{i,t-p}(\mathbf{w}, \tilde{\mathbf{w}}; p) = \hat{\tau}_{i,t}(\mathbf{w}, \tilde{\mathbf{w}}; p) - \tau_{i,t}(\mathbf{w}, \tilde{\mathbf{w}}; p)$ be the estimation error. Now

$$\begin{aligned} u_{i,t-p}(\mathbf{w}, \tilde{\mathbf{w}}; p) = Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big)\big(Z_{i,t-p:t}(\mathbf{w}) - 1\big) \\ - Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big)\big(Z_{i,t-p:t}(\tilde{\mathbf{w}}) - 1\big). \end{aligned}$$

Hence the conditional expectation is zero by Lemma A.1. Then

$$\begin{aligned} &\text{Var}\big(u_{i,t-p}(\mathbf{w}, \tilde{\mathbf{w}}; p) | \mathcal{F}_{i,t-p-1}\big) \\ &= Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big)^2 \text{Var}\big(Z_{i,t-p:t}(\mathbf{w}) | \mathcal{F}_{i,t-p-1}\big) \\ &\quad + Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big)^2 \text{Var}\big(Z_{i,t-p:t}(\tilde{\mathbf{w}}) | \mathcal{F}_{i,t-p-1}\big) \\ &\quad - 2Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big) \text{Cov}(Z_{i,t-p:t}(\tilde{\mathbf{w}}), Z_{i,t-p:t}(\tilde{\mathbf{w}} | \mathcal{F}_{i,t-p-1}) \\ &= Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big)^2 p_{i,t-p}(\mathbf{w})^{-1}\big(1 - p_{i,t-p}(\mathbf{w})\big) \\ &\quad + Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big)^2 p_{i,t-p}(\tilde{\mathbf{w}})^{-1}\big(1 - p_{i,t-p}(\tilde{\mathbf{w}})\big) \\ &\quad - 2Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big). \end{aligned}$$

Simplifying gives the result on the variance of the estimation error. Then

$$\begin{aligned} &\text{Cov}\big(u_{i,t-p}(\mathbf{w}, \tilde{\mathbf{w}}; p), u_{i,t-p}(\bar{\mathbf{w}}, \hat{\mathbf{w}}; p) | \mathcal{F}_{i,t-p-1}\big) \\ &= Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \bar{\mathbf{w}}\big) \text{Cov}\big(Z_{i,t-p:t}(\mathbf{w}), Z_{i,t-p:t}(\bar{\mathbf{w}}) | \mathcal{F}_{i,t-p-1}\big) \\ &\quad - Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \hat{\mathbf{w}}\big) \text{Cov}\big(Z_{i,t-p:t}(\mathbf{w}), Z_{i,t-p:t}(\hat{\mathbf{w}}) | \mathcal{F}_{i,t-p-1}\big) \\ &\quad - Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \bar{\mathbf{w}}\big) \text{Cov}\big(Z_{i,t-p:t}(\tilde{\mathbf{w}}), Z_{i,t-p:t}(\bar{\mathbf{w}}) | \mathcal{F}_{i,t-p-1}\big), \\ &Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \hat{\mathbf{w}}\big) \text{Cov}\big(Z_{i,t-p:t}(\tilde{\mathbf{w}}), Z_{i,t-p:t}(\hat{\mathbf{w}}) | \mathcal{F}_{i,t-p-1}\big) \\ &= -Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \bar{\mathbf{w}}\big) + Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \hat{\mathbf{w}}\big) \\ &\quad + Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \bar{\mathbf{w}}\big) - Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big) Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \hat{\mathbf{w}}\big) \end{aligned}$$

Finally, conditional independence of the errors follows due to the individualistic assignment of treatments.

Proof of Proposition 3.1

The proof of this result is analogous to the proof of Theorem 3.1. We state the analogue of Lemma A.1 for completeness.

LEMMA A.2. *Assume a potential outcome panel with an assignment mechanism that is individualistic* (*Definition* 3) *and probabilistic* (*Assumption* 2). *Define, for any* $\mathbf{w} \in \mathcal{W}^{(p+1)}$, *the random function* $V_{i,t-p:t}(\mathbf{w}) := p_{i,t-p}(\mathbf{w})^{-2} \mathbb{1}\{W_{i,t-p:t} = \mathbf{w}\}$. *Then,*

over the assignment mechanism, $\mathbb{E}(V_{i,t-p:t}(\mathbf{w})|\mathcal{F}_{i,t-p-1}) = p_{i,t-p}(\mathbf{w})^{-1}$ *and* $\mathrm{Var}(V_{i,t-p:t}(\mathbf{w})|\mathcal{F}_{i,t-p-1}) = p_{i,t-p}(\mathbf{w})^{-3}(1-p_{i,t-p}(\mathbf{w}))$, *and* $\mathrm{Cov}(V_{i,t-p:t}(\mathbf{w}), V_{i,t-p:t}(\tilde{\mathbf{w}})|\mathcal{F}_{i,t-p-1}) = p_{i,t-p}(\mathbf{w})^{-1}p_{i,t-p}(\tilde{\mathbf{w}})^{-1}$ *for all* $\mathbf{w} \neq \tilde{\mathbf{w}}$. *Moreover,* $V_{i,t-p:t}(\mathbf{w})$ *and* $V_{j,t-p:t}(\mathbf{w})$ *are, conditioning on* $\mathcal{F}_{1:N,t-p-1}$, *independent for* $i \neq j$.

For any $\mathbf{w}, \tilde{\mathbf{w}} \in \mathcal{W}^{(p+1)}$, let $v_{i,t-p}(\mathbf{w}, \tilde{\mathbf{w}}; p) = \hat{\gamma}^2_{i,t}(\mathbf{w}, \tilde{\mathbf{w}}; p) - \gamma^2_{i,t}(\mathbf{w}, \tilde{\mathbf{w}}; p)$ be the estimation error. Now

$$\begin{aligned} v_{i,t-p}(\mathbf{w}, \tilde{\mathbf{w}}; p) = {} & Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \mathbf{w}\big)^2\big(V_{i,t-p:t}(\mathbf{w}) - p_{i,t-p}(\mathbf{w})^{-1}\big) \\ & + Y_{i,t}\big(w^{\text{obs}}_{i,1:t-p-1}, \tilde{\mathbf{w}}\big)^2\big(V_{i,t-p:t}(\tilde{\mathbf{w}}) - \tilde{p}_{i,t-p}(\mathbf{w})^{-1}\big). \end{aligned}$$

Therefore, the conditional expectation is zero by Lemma A.2. The conditional independence of the errors follows due to the individualistic assignment of the treatments.

Proof of Theorem 3.2

Only the third results requires a new proof. The first result is a reinterpretation of the classic cross-sectional result using a triangular array central limit theorem, for the usual Lindeberg condition must hold due to the bounded potential outcomes and the treatments being probabilistic. The second result follows from results in Bojinov and Shephard (2019), who use a martingale difference array central limit theorem.

The third result, which holds for NT going to infinity, can be split into three parts. For NT to go to infinity, we must have either: (i) T goes to infinity with N finite, (ii) N goes to infinity with T finite, or (iii) both N and T go to infinity. In the case (i), we apply the martingale difference CLT but now we have preaveraged the cross-sectional errors over the N terms for each time period. The preaverage is still a martingale difference, so the technology is the same. In the case (ii), we preaverage the time aspect. Then we are back to a standard triangular array CLT. As we have both (i) and (ii), then (iii) must hold.

Proof of Proposition 3.2

The unbiasedness statements follow directly from Proposition 3.1. The proofs of the consistency statements are analogous to the proof of Theorem 3.2. The first result follows from an application of the triangular array law of law of large numbers, which may be applied due to the bounded potential outcomes and the treatments being probabilistic. The second statement follows from an application of a martingale difference sequence law of large numbers (Theorem 2.13 in Hall and Heyde (1980)). The third statement can be again proved in three cases: (i) T goes to infinity with N finite, (ii) N goes to infinity with T finite, or (iii) both N and T go to infinity as in the proof of Theorem 3.2 and applying the appropriate law of large numbers.

Proof of Proposition 4.1

Begin by writing the observed outcomes as

$$Y_{i,t} = Y_{i,t}(\mathbf{0}) + \sum_{s=1}^{t} \beta_{i,t,t-s} W_{i,s}.$$

Similarly, write $\bar{Y}_{i\cdot} = \bar{Y}_{i\cdot}(\mathbf{0}) + \overline{\beta W}_{i\cdot}$, where $\overline{\beta W}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{t}\beta_{i,t,t-s}W_{i,s}$. The transformed outcome can be then written as

$$\breve{Y}_{i,t} = \sum_{s=1}^{t}\beta_{i,t,t-s}W_{i,s} - \overline{\beta W}_{i\cdot} + \breve{Y}_{i,t}(\mathbf{0}).$$

Consider the numerator of the unit fixed effects estimator. Substituting in, we arrive at

$$\begin{aligned}
\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\breve{Y}_{i,t}\breve{W}_{i,t} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\beta_{i,t,0}W_{i,t}\breve{W}_{i,t} \\
&\quad + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\sum_{s=1}^{t-1}\beta_{i,t,t-s}W_{i,s}\breve{W}_{i,t}\right) \\
&\quad + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\breve{Y}_{i,t}(\mathbf{0})\breve{W}_{i,t} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\beta_{i,t,0}W_{i,t}\breve{W}_{i,t}\right) \\
&\quad + \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left(\frac{1}{N}\sum_{i=1}^{N}\beta_{i,t,t-s}W_{i,s}\breve{W}_{i,t}\right) \\
&\quad + \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\breve{Y}_{i,t}(\mathbf{0})\breve{W}_{i,t}\right).
\end{aligned}$$

Therefore, for fixed T as $N \to \infty$,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\beta_{i,t,0}W_{i,t}\breve{W}_{i,t}\right) &\xrightarrow{p} \frac{1}{T}\sum_{t=1}^{T}\breve{\kappa}_{W,\beta,t,t}, \\
\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\left(\frac{1}{N}\sum_{i=1}^{N}\beta_{i,t,t-s}W_{i,s}\breve{W}_{i,t}\right) &\xrightarrow{p} \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\breve{\kappa}_{W,\beta,t,s}, \\
\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\breve{Y}_{i,t}(\mathbf{0})\breve{W}_{i}\right) &\xrightarrow{p} \frac{1}{T}\sum_{t=1}^{T}\breve{\delta}_{t}.
\end{aligned}$$

Similarly, the denominator converges to $\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\breve{W}_{i,t}^2 \xrightarrow{p} \frac{1}{T}\sum_{t=1}^{T}\breve{\sigma}_{W,t}^2$. The result then follows by Slutsky.

Proof of Proposition 4.2

Begin by writing

$$Y_{i,t} = Y_{i,t}(\mathbf{0}) + \sum_{s=1}^{t}\beta_{i,t,t-s}W_{i,s}.$$

Then $\bar{Y}_{\cdot t} = \bar{Y}_{\cdot t}(\mathbf{0}) + \overline{\beta W}_{\cdot t}$, $\bar{Y}_{i\cdot} = \bar{Y}_{i\cdot}(\mathbf{0}) + \overline{\beta W}_{i\cdot}$ and $\bar{Y} = \bar{Y}(\mathbf{0}) + \overline{\beta W}$. Therefore,

$$\dot{\breve{Y}}_{i,t} = \dot{\breve{Y}}_{i,t}(\mathbf{0}) + \left(\sum_{s=1}^{t} \beta_{i,t,t-s} W_{i,s} - \overline{\beta W}\right) - (\overline{\beta W}_{\cdot t} - \overline{\beta W}) - (\overline{\beta W}_{i\cdot} - \overline{\beta W}).$$

Consider the numerator of the unit fixed effects estimator. Substituting in,

$$\begin{aligned}
\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \dot{\breve{Y}}_{i,t}\dot{\breve{W}}_{i,t} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \beta_{i,t,0} W_{i,t}\dot{\breve{W}}_{i,t} \\
&\quad + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{t-1} \beta_{i,t,t-s} W_{i,s}\dot{\breve{W}}_{i,t} \\
&\quad + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} \dot{\breve{Y}}_{i,t}(\mathbf{0})\dot{\breve{W}}_{i,t} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} \beta_{i,t,0} W_{i,t}\dot{\breve{W}}_{i,t}\right) \\
&\quad + \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{t-1} \beta_{i,t,t-s} W_{i,s}\dot{\breve{W}}_{i,t}\right) \\
&\quad + \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} \dot{\breve{Y}}_{i,t}(\mathbf{0})\dot{\breve{W}}_{i,t}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N} \beta_{i,t,0} W_{i,t}\dot{\breve{W}}_{i,t} &\xrightarrow{p} \dot{\breve{\kappa}}_{W,\beta,t,t}, \\
\frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{t-1} \beta_{i,t,t-s} W_{i,s}\dot{\breve{W}}_{i,t} &\xrightarrow{p} \sum_{s=1}^{t-1} \dot{\breve{\kappa}}_{W,\beta,t,s}, \\
\frac{1}{N}\sum_{i=1}^{N} \dot{\breve{Y}}_{i,t}(\mathbf{0})\dot{\breve{W}}_{i,t} &\xrightarrow{p} \dot{\breve{\delta}}_{t}.
\end{aligned}$$

A similar argument applies to the denominator and the result follows.

Appendix B: Additional theoretical results

B.1 *Prediction decomposition of the adapted propensity score*

Recall the definition of the adapted propensity score in Section 3:

$$p_{i,t-p}(\mathbf{w}) := \Pr\big(W_{i,t-p:t} = \mathbf{w} | W_{i,1:t-p-1},\, Y_{i,1:t}(W_{i,1:t-p-1}, \mathbf{w})\big).$$

The adapted propensity score can be decomposed using individualistic assignment (Definition 3) and the prediction decomposition.

LEMMA B.1. *For a potential outcome panel satisfying individualistic assignment* (*Definition* 3) *and any* $\mathbf{w} \in \mathcal{W}^{(p+1)}$, *the adapted propensity score can be factorized as*

$$\begin{aligned} p_{i,t-p}(\mathbf{w}) = \Pr\big(&W_{i,t-p} = w_1 | W_{i,1:t-p-1}, Y_{i,1:t-p-1}(W_{i,1:t-p-1})\big) \\ &\times \prod_{s=1}^{p} \Pr\big(W_{i,t-p+s} = w_{s+1} | W_{i,1:t-p-1}, W_{i,t-p:t-p+s-1} = \mathbf{w}_{1:s}, \\ &Y_{i,1:t-p+s-1}(W_{i,1:t-p-1}, \mathbf{w}_{1:s})\big). \end{aligned}$$

PROOF. Use the prediction decomposition for assignments, given all outcomes,

$$\begin{aligned} p_{i,t-p}(\mathbf{w}) = \Pr\big(&W_{i,t-p} = w_1 | W_{i,1:t-p-1}, Y_{i,1:t}(W_{i,1:t-p-1}, \mathbf{w})\big) \\ &\times \prod_{s=1}^{p} \Pr\big(W_{i,t-p+s} = w_{s+1} | W_{i,1:t-p-1}, W_{i,t-p:t-p+s-1} = \mathbf{w}_{1:s}, \\ &Y_{i,1:t}(W_{i,1:t-p-1}, \mathbf{w})\big). \end{aligned}$$

and then simplify using the individualistic assignment of treatments. □

Even though the assignment mechanism is known, we only observe the outcomes along the realized assignment path $Y_{i,1:t}(w^{\text{obs}}_{i,1:t})$, and so it is not possible to use Lemma B.1 to compute $p_{i,t-p}(\mathbf{w})$ for all assignment path. We can, however, compute the adapted propensity score along the observed assignment path, $p_{i,t-p}(w^{\text{obs}}_{i,t-p:t})$, since the associated outcomes are observed.

B.2 *Estimation as a repeated cross-section*

Denote $\dot{Y}_{1:N,t} = (\dot{Y}_{1,t}, \ldots, \dot{Y}_{N,t})'$, $\dot{W}_{i,1:t} = (W_{i,t} - \bar{W}_{\cdot t}, W_{i,t-1} - \bar{W}_{\cdot t-1}, \ldots, W_{i,1} - \bar{W}_{\cdot 1})'$ and $\dot{W}_{1:N,t} = (\dot{W}_{1,1:t}, \ldots, \dot{W}_{N,1:t})'$. The least squares coefficient in the regression of $\dot{Y}_{1:N,t}$ on $\dot{W}_{1:N,t}$ is $\hat{\boldsymbol{\beta}}_{1:N,t} = (\dot{W}'_{1:N,t}\dot{W}_{1:N,t})^{-1}\dot{W}'_{1:N,t}\dot{Y}_{1:N,t}$. Proposition B.1 derives the finite population limiting distribution of $\hat{\boldsymbol{\beta}}_{1:N,t}$ as the number of units grows large.

PROPOSITION B.1. *Assume a potential outcome panel and consider the "control" only path, for* $0 \in \mathcal{W}$ *let* $\tilde{w}_{i,1:t} = \mathbf{0}$. *Let* $\dot{\mu}_{i,t}$ *be the* $t \times 1$ *vector whose* u*th element is* $E[\dot{W}_{i,t-(u-1)} | \mathcal{F}_{1:N,0,T}]$ *and* $\Omega_{i,t}$ *be the* $t \times t$ *matrix whose* u, v*th element is* $\text{Cov}(\dot{W}_{i,t-(u-1)}, \dot{W}_{i,t-(v-1)} | \mathcal{F}_{1:N,0,T})$. *Additionally, assume that*:

1. *The potential outcome panel is linear* (*Definitions* 7) *and homogeneous with* $\boldsymbol{\beta}_{it} \equiv \boldsymbol{\beta}_t = (\beta_{t,0}, \ldots, \beta_{t,t-1})$ *for all* t.
2. $W_{i,1:t}$ *is an individualistic stochastic assignment path and, over the randomization distribution,* $\text{Var}(W_{i,t} | \mathcal{F}_{1:N,0,T}) = \sigma^2_{W,i,t} < \infty$ *for each* $i \in [N]$, $t \in [T]$.

3. *As* $N \to \infty$,

   (a) *Nonstochastically,* $N^{-1}\sum_{i=1}^{N}\Omega_{i,t} \to \Gamma_{2,t}$, *where* $\Gamma_{2,t}$ *is positive definite.*

   (b) $N^{-1/2}\sum_{i=1}^{N}(\dot{W}_{i,1:t} - \dot{\mu}_{i,t})\dot{Y}_{i,t}(\mathbf{0})|\mathcal{F}_{1:N,0,T} \xrightarrow{d} N(0,\Gamma_{1,t})$.

   (c) *Nonstochastically,* $N^{-1}\sum_{i=1}^{N}\dot{Y}_{i,t}(\mathbf{0})\dot{\mu}_{i,t} \to \dot{\delta}_t$.

Then, over the randomization distribution, as $N \to \infty$,

$$\sqrt{N}\big(\hat{\boldsymbol{\beta}}_{1:N,t} - \boldsymbol{\beta}_t - \Gamma_{2,t}^{-1}\dot{\delta}_t\big)|\mathcal{F}_{1:N,0,T} \xrightarrow{d} N\big(0, \Gamma_{2,t}^{-1}\Gamma_{1,t}\Gamma_{2,t}^{-1}\big).$$

PROOF. Under linear potential outcomes,

$$Y_{i,t}(W_{i,1:t}) - Y_{i,t}(\tilde{W}_{i,1:t}) = \sum_{s=0}^{t-1}\beta_{i,t,s}(W_{i,t-s} - \tilde{W}_{i,t-s}).$$

Focus on the counterfactual $\tilde{W}_{i,1:t} = \mathbf{0}$, then

$$Y_{i,t} = Y_{i,t}(W_{i,1:t}) = \bar{Y}_{\cdot t}(\mathbf{0}) + \sum_{s=0}^{t-1}\beta_{i,t,s}W_{i,t-s} + \dot{Y}_{i,t}(\mathbf{0}).$$

Therefore, the within-period transformed outcome equals

$$\dot{Y}_{i,t} = Y_{i,t} - \bar{Y}_{\cdot t} = \sum_{s=0}^{t-1}\left\{\beta_{i,t,s}W_{i,t-s} - \frac{1}{N}\sum_{j=1}^{N}\beta_{j,t,s}W_{j,t-s}\right\} + \dot{Y}_{i,t}(\mathbf{0}).$$

Further imposing homogeneity, it simplifies to

$$\dot{Y}_{i,t} = \sum_{s=0}^{t-1}\left\{\beta_{t,s}(W_{i,t-s} - \frac{1}{N}\sum_{j=1}^{N}W_{j,t-s}\right\} + \dot{Y}_{i,t}(\mathbf{0}).$$

Stacking everything across units, this becomes $\dot{Y}_{1:N,t} = \dot{W}_{1:N,t}\beta_t + \dot{Y}_{1:N,t}(\mathbf{0})$, and so the linear projection coefficient is given by

$$\hat{\beta}_t = \big(\dot{W}'_{1:N,t}\dot{W}_{1:N,t}\big)^{-1}\dot{W}'_{1:N,t}\dot{Y}_{1:N,t} = \beta_t + \big(\dot{W}'_{1:N,t}\dot{W}_{1:N,t}\big)^{-1}\dot{W}'_{1:N,t}\dot{Y}_{1:N,t}(\mathbf{0}).$$

The important unusual point here is that $\dot{Y}_{1:N,t}(\mathbf{0})$ is nonstochastic and that $\dot{W}_{1:N,t}$ is random, exactly the opposite of the case often discussed in the statistical analysis of linear regression. Now

$$\frac{1}{N}\dot{W}'_{1:N,t}\dot{W}_{1:N,t} = \frac{1}{N}\sum_{i=1}^{N}\dot{W}_{i,1:t}\dot{W}'_{i,1:t},$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\dot{W}_{i,1:t}\dot{Y}_{i,t}(\mathbf{0}) = \frac{1}{N}\sum_{i=1}^{N}(\dot{W}_{i,1:t} - \dot{\mu}_{i,t})\dot{Y}_{i,t}(\mathbf{0}) + \frac{1}{N}\sum_{i=1}^{N}\mu_{i,t}\dot{Y}_{i,t}(\mathbf{0}).$$

Then, under the assumption of individualistic assignment (Definition 3),

$$\frac{1}{N}\sum_{i=1}^{N}\dot{W}_{i,1:t}\dot{W}'_{i,1:t}|\mathcal{F}_{1:N,0,T} \xrightarrow{p} \Gamma_{2,t},$$

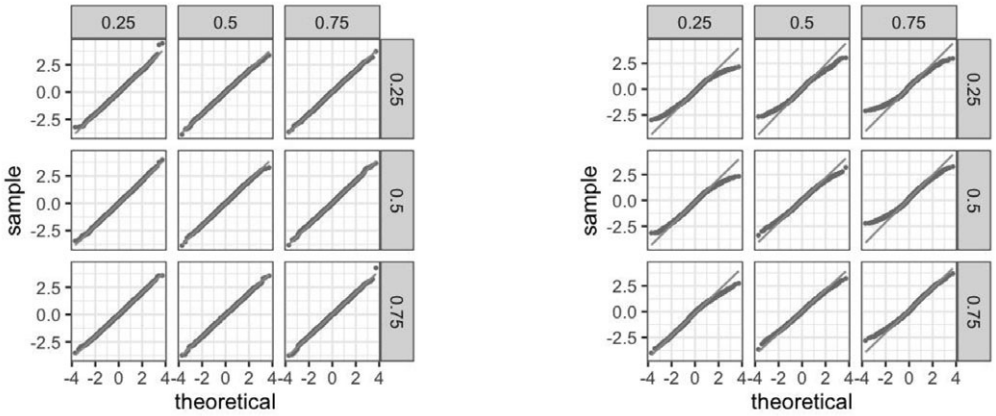
recalling $\dot{Y}_{i,t}(\mathbf{0})$ is nonstochastic and applying Assumptions 3(b) and 3(c), then Slutsky's theorem delivers the result. □

Appendix C: Additional simulation results

C.1 *Additional simulations for the estimator of the total average dynamic causal effects*

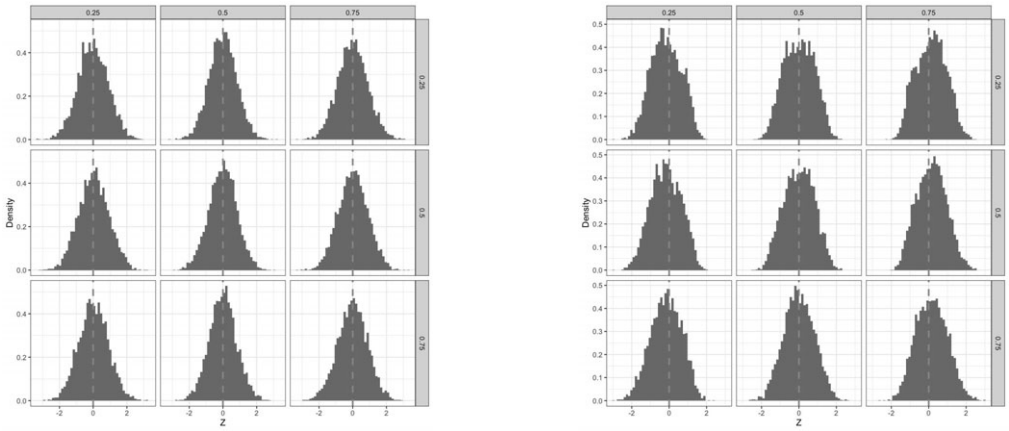
Quantile-quantile plot for the normal approximation Figure A1 provides quantile-quantile plots of the simulated randomization distribution for the estimator $\hat{\bar{\tau}}(1,0;0)$ presented in Section 5 of the main text.

*Simulation results for the estimator of the lag-*1 *total weighted average dynamic causal effect,* $\bar{\tau}^{\dagger}(1,0;1)$ We now present simulation results that analyze the properties of our estimator for the lag-1 total weighted average dynamic causal effect, $\hat{\bar{\tau}}^{\dagger}(1,0;1)$. We choose the weights to $a_{\mathbf{v}}$ to place equal weight on the future treatment paths. Figure A2 plots the simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}(1,0;1)$ and Figure A3 plots the associated



(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$, $T = 10$ (b) $\epsilon_{i,t} \sim Cauchy$, $N = 500$, $T = 100$

Figure A1. Quantile-quantile plots for the simulated randomization distribution for $\hat{\bar{\tau}}(1,0;0)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The quantile-quantile plots compare the quantiles of the simulated randomization distribution (y-axis) against the quantiles of a standard normal random variable (x-axis). The 45-degree line by the solid line. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the quantile-quantile plots for simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$, $T = 10$. Panel (b) plots the quantile-quantile plots simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim Cauchy$ and $N = 500$, $T = 100$. Results are computed over 5000 simulations. See Section 5 of the main text for further details.



(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$, $T = 10$ (b) $\epsilon_{i,t} \sim Cauchy$, $N = 500$, $T = 100$

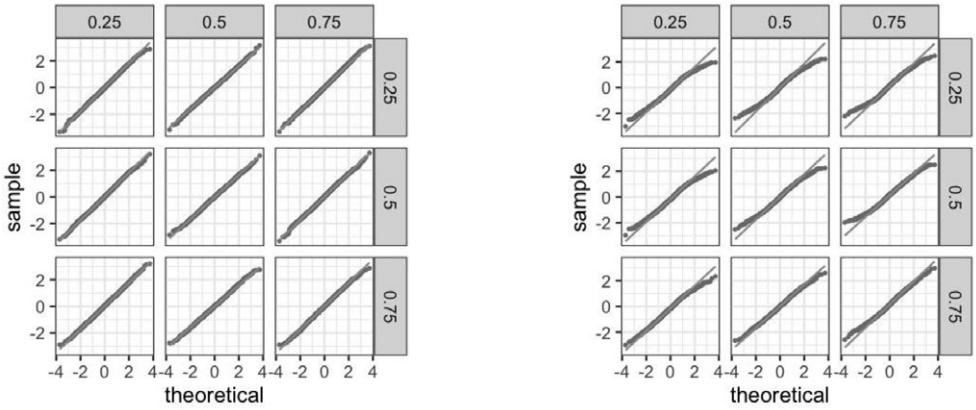
FIGURE A2. Simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}(1,0;1)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$, $T = 10$. Panel (b) plots the simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim Cauchy$ and $N = 500$, $T = 10$. Results are computed over 5000 simulations. See Section 5 of the main text for further details.

quantile-quantile plot. We observe that the normal approximation remains accurate for lagged dynamic causal effects.

C.2 *Simulations for the estimator of the time-t average dynamic causal effects*

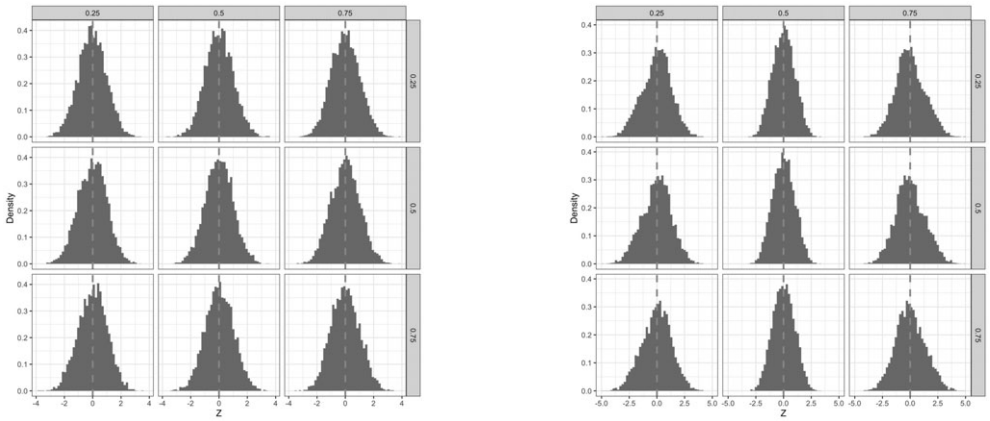
We present simulation results for our estimator of the time-t average dynamic causal effect, $\hat{\bar{\tau}}_{\cdot,t}(1,0;0)$, with $N = 100$ units when the potential outcomes are generated with normally distributed errors and $N = 50{,}000$ with Cauchy distributed errors.

Normal approximations and size control Figure A4 plots the randomization distribution for the estimator of the contemporaneous time-t average dynamic causal effect, $\hat{\bar{\tau}}_{\cdot t}(1,0;0)$, under the null hypothesis of $\beta = 0$ for different combinations of the parameter $\phi \in \{0.25, 0.5, 0.75\}$ and treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. When the errors $\epsilon_{i,t}$ are normally distributed, the randomization distribution quickly converges to a normal distribution—the normal approximation is accurate when there are only $N = 100$ units in the experiment. As expected, when the errors are Cauchy distributed, the number of units must be quite large for the randomization distribution to become approximately normal. There is little difference in the results across the values of ϕ and $p(w)$. Figure A5 provides quantile-quantile plots of the simulated randomization distributions to further illustrate the quality of the normal approximations. Testing based on the normal asymptotic approximation controls size effectively, staying close to the nominal 5% level (the exact rejection rates for the null hypothesis, $H_0 : \bar{\tau}_{\cdot t}(1,0;0) = 0$ are reported in Table A1).



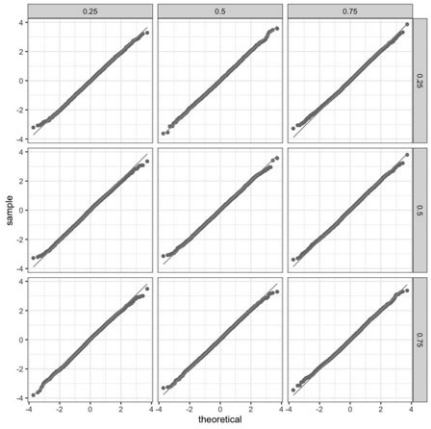
(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100, T = 10$ (b) $\epsilon_{i,t} \sim Cauchy$, $N = 500, T = 100$

FIGURE A3. Quantile-quantile plots for the simulated randomization distribution for $\hat{\tilde{\tau}}^{\dagger}(1,0;1)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The quantile-quantile plots compare the quantiles of the simulated randomization distribution (y-axis) against the quantiles of a standard normal random variable (x-axis). The 45-degree line is plotted by the solid line. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the quantile-quantile plots for simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 1000$. Panel (b) plots the quantile-quantile plots simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $T = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details.

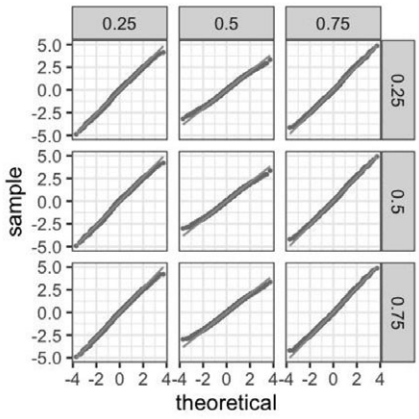


(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$ (b) $\epsilon_{i,t} \sim Cauchy$, $N = 50,000$

FIGURE A4. Simulated randomization distribution for $\hat{\tilde{\tau}}_{\cdot t}(1,0;0)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$ and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$. Panel (b) plots the simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $N = 50{,}000$. Results are computed over 5000 iterations. See Section 5 of the main text for further details on the simulation design.



(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$



(b) $\epsilon_{i,t} \sim Cauchy$, $N = 50,000$

FIGURE A5. Quantile-quantile plots for the simulated randomization distribution for $\hat{\bar{\tau}}_{\cdot t}(1,0;0)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The quantile-quantile plots compare the quantiles of the simulated randomization distribution (y-axis) against the quantiles of a standard normal random variable (x-axis). The 45-degree line is plotted by the solid line. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the quantile-quantile plots for simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$. Panel (b) plots the quantile-quantile plots simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $N = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

Rejection rates Figure A6 plots rejection rate curves against the null hypotheses as the parameter β varies for different choices of the parameter ϕ and treatment probability $p(w)$ in simulations with $N = 100$ units. For $p = 0$, the rejection rate against $H_0 : \bar{\tau}_{\cdot t}(1,0;0) = 0$ quickly converges to one as β moves away from zero across a range of simulations. This is encouraging as it indicates that the conservative variance bound

TABLE A1. Null rejection rate for the test of the null hypothesis $H_0 : \bar{\tau}_{\cdot t}(1,0;0) = 0$ based upon the normal asymptotic approximation to the randomization distribution of $\hat{\bar{\tau}}_{\cdot t}(1,0;0)$. Panel (a) reports the null rejection probabilities in simulations with $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$. Panel (b) reports the null rejection probabilities in simulations with $\epsilon_{i,t} \sim$ *Cauchy* and $N = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

		$p(w)$			$p(w)$		
		0.25	0.5	0.75	0.25	0.5	0.75
		(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$			(b) $\epsilon_{i,t} \sim$ *Cauchy*, $N = 50{,}000$		
ϕ	0.25	0.046	0.048	0.048	0.031	0.031	0.034
	0.5	0.049	0.049	0.050	0.048	0.039	0.043
	0.75	0.050	0.049	0.045	0.052	0.047	0.057

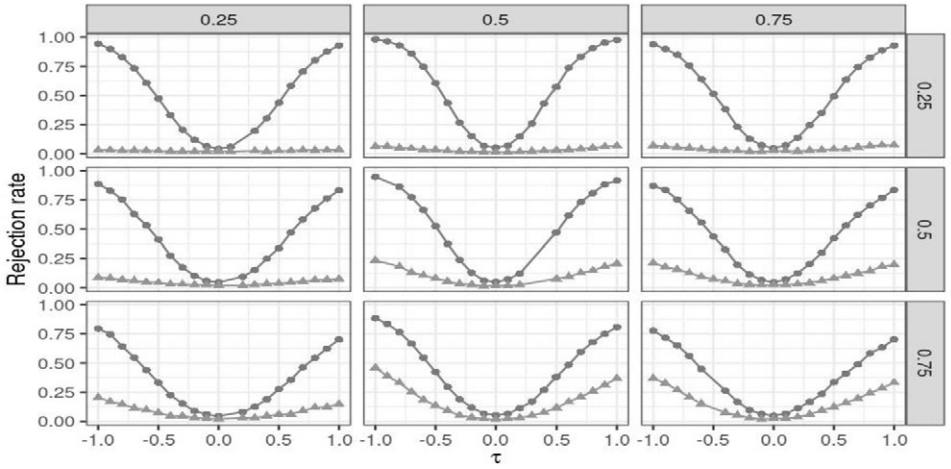


FIGURE A6. Rejection probabilities for a test of the null hypothesis $H_0 : \bar{\tau}_{\cdot,t}(1,0;0) = 0$ and $H_0 : \bar{\tau}^{\dagger}_{\cdot,t}(1,0;1) = 0$ as the parameter β varies under different choices of the parameter ϕ and treatment probability $p(w)$. The rejection rate curve against $H_0 : \bar{\tau}_{\cdot,t}(1,0;0) = 0$ is plotted by the circles and the rejection rate curve against $H_0 : \bar{\tau}^{\dagger}_{\cdot,t}(1,0;1) = 0$ is plotted by the triangles. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. The simulations are conducted with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$. Results are averaged over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

still leads to informative tests. However, when $p = 1$, the persistence of the causal effects ϕ has an important effect on the power of our tests. In particular, when $\phi = 0.25$, the rejection rate against $H_0 : \bar{\tau}^{\dagger}_{\cdot,t}(1,0;1) = 0$ is quite low for all values of β – lower values of ϕ imply less persistence in the causal effects across periods. When $\phi = 0.75$, there is substantial persistence across periods and observe that the rejection rate curves improve for $p = 1$. Additionally, Figure A7 shows the same power plots for $N = 1000$ units. We again observe that power is relatively low for low values of ϕ, but when $\phi = 0.75$, the rejection rate curves for $p = 0, 1$ appear similar. This suggests that detecting dynamic causal effects requires larger sample sizes.

*Simulation results for the estimator of the lag-*1*, time-t weighted average dynamic causal effect,* $\bar{\tau}^{\dagger}_{\cdot,t}(1,0;1)$ We now present simulation results that analyze the properties of our estimator for the lag-1 total weighted average dynamic causal effect, $\hat{\bar{\tau}}^{\dagger}_{\cdot,t}(1,0;1)$. We choose the weights to $a_{\mathbf{v}}$ to place equal weight on the future treatment paths. Figure A8 plots the simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}_{\cdot,t}(1,0;1)$ and Figure A9 plots the associated quantile-quantile plot. We observe that the normal approximation remains accurate for lagged dynamic causal effects.

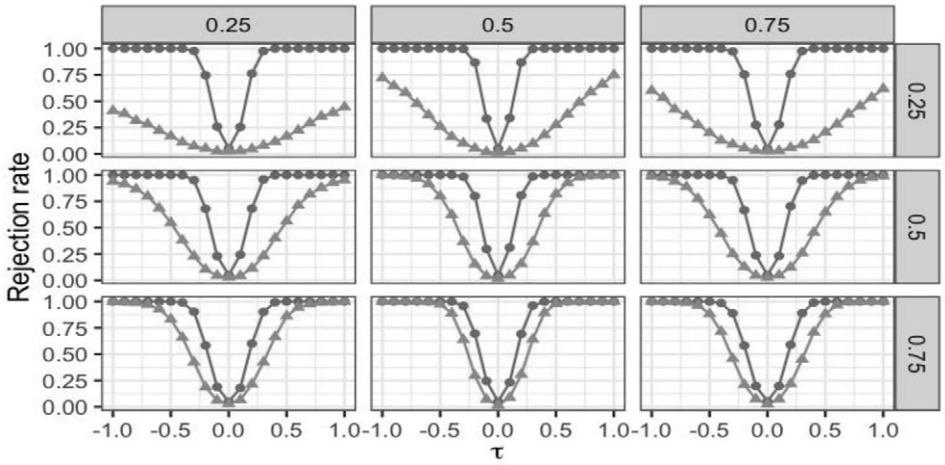
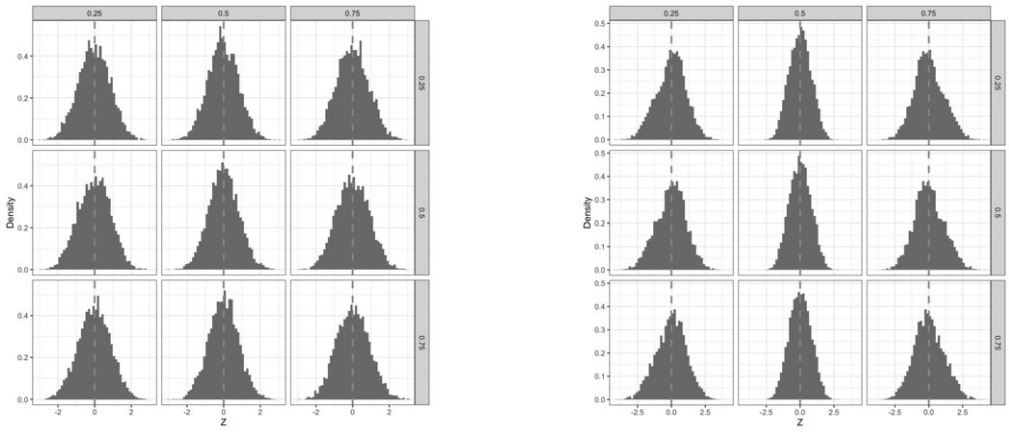
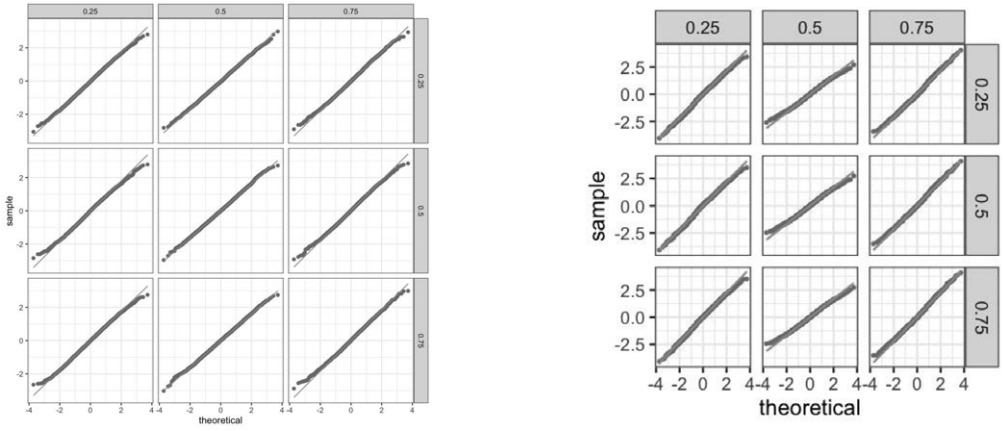


FIGURE A7. Rejection probabilities for a test of the null hypothesis $H_0 : \bar{\tau}_{\cdot t}(1,0;0) = 0$ and $H_0 : \bar{\tau}^{\dagger}_{\cdot t}(1,0;1) = 0$ as the parameter β varies under different choices of the parameter ϕ and treatment probability $p(w)$. The rejection rate curve against $H_0 : \bar{\tau}_{\cdot t}(1,0;0) = 0$ is plotted by the circles and the rejection rate curve against $H_0 : \bar{\tau}^{\dagger}_{\cdot t}(1,0;1) = 0$ is plotted by the triangles. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. The simulations are conducted with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 1000$. Results are averaged over 5000 simulations. See Section 5 of the main text for further details on the simulation design.



(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$ (b) $\epsilon_{i,t} \sim Cauchy$, $N = 50,000$

FIGURE A8. Simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}_{\cdot t}(1,0;1)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 100$. Panel (b) plots the simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $N = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.



(a) $\epsilon_{i,t} \sim N(0,1)$, $N = 100$ (b) $\epsilon_{i,t} \sim Cauchy$, $N = 50,000$

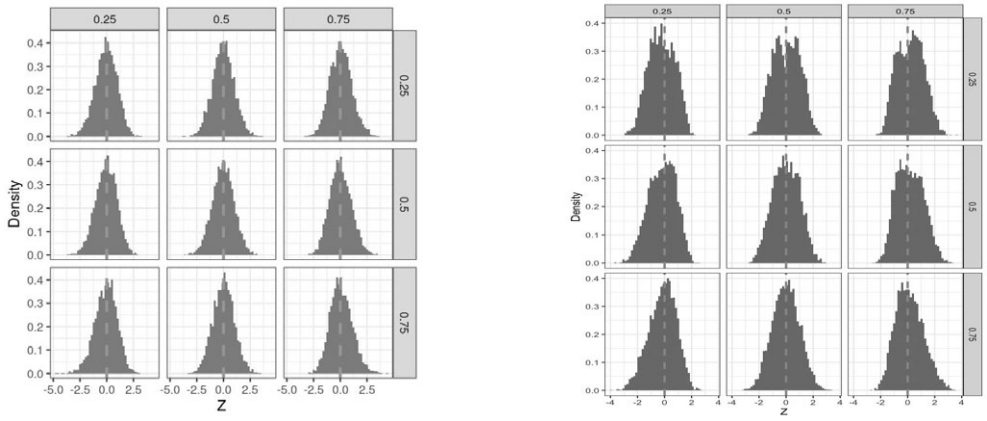
FIGURE A9. Quantile-quantile plots for the simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}_{\cdot,t}(1,0;1)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The quantile-quantile plots compare the quantiles of the simulated randomization distribution (y-axis) against the quantiles of a standard normal random variable (x-axis). The 45-degree line is plotted by the solid line. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the quantile-quantile plots for simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $N = 1000$. Panel (b) plots the quantile-quantile plots simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim Cauchy$ and $N = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

C.3 *Simulations for the estimator of the unit-i average dynamic causal effects*

We present simulation results for our estimator of the unit-i average dynamic causal effect, $\hat{\bar{\tau}}_{i,\cdot}(1,0;0)$, with $T = 100$ time periods when the potential outcomes are generated with normally distributed errors and $T = 50{,}000$ with Cauchy distributed errors.

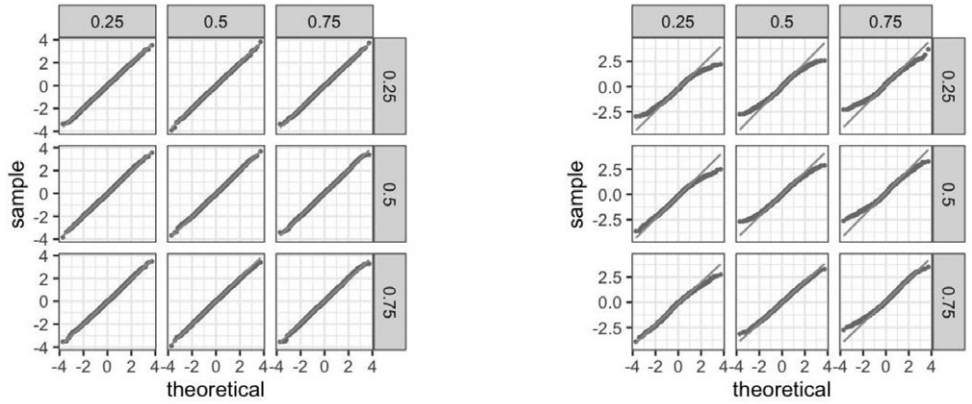
Normal approximations and size control Figure A10 plots the randomization distribution for $\hat{\bar{\tau}}_{i\cdot}(1,0;0)$. We see a similar pattern as before—when the errors are normally distributed, the randomization distribution converges quickly to a normal distribution, but it takes longer to do so when the errors are heavy-tailed. Figure A11 provides quantile-quantile plots of the simulation randomization distributions to further illustrate the quality of the normal approximations. The null rejection rates for the hypothesis, $H_0 : \bar{\tau}_{i,\cdot}(1,0;0) = 0$ are reported in Table A2 and, again, the test controls size well across a wide range of parameters.

Rejection rates Next, we investigate the rejection rate of the statistical test based on the normal asymptotic approximation for $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;0) = 0$ and $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;1) = 0$, plotting the rejection rates in Figure A12. For $p = 0$, Once again, we observe that the rejection rate against $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;0) = 0$ has good power properties across a range of simulations. However, once again for $p = 1$, our conservative test has low power and the persistence of the causal effects ϕ has an important effect on the power of our tests. Additionally,



(a) $\epsilon_{i,t} \sim N(0,1)$, $T = 100$ (b) $\epsilon_{i,t} \sim Cauchy$, $T = 50,000$

FIGURE A10. Simulated randomization distribution for $\hat{\bar{\tau}}_{i\cdot}(1,0;0)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 100$. Panel (b) plots the simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $T = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.



(a) $\epsilon_{i,t} \sim N(0,1)$, $T = 100$ (b) $\epsilon_{i,t} \sim Cauchy$, $T = 50,000$

FIGURE A11. Quantile-quantile plots for the simulated randomization distribution for $\hat{\bar{\tau}}_{i\cdot}(1,0;0)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The quantile-quantile plots compare the quantiles of the simulated randomization distribution (y-axis) against the quantiles of a standard normal random variable (x-axis). The 45-degree line is plotted by the solid line. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the quantile-quantile plots for simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 100$. Panel (b) plots the quantile-quantile plots simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $T = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

TABLE A2. Null rejection rate for the test of the null hypothesis $H_0 : \bar{\tau}_{i\cdot}(1,0;0) = 0$ based upon the normal asymptotic approximation to the randomization distribution of $\hat{\bar{\tau}}_{i\cdot}(1,0;0)$. Panel (a) reports the null rejection probabilities in simulations with $\epsilon_{i,t} \sim N(0,1)$ and $T = 100$. Panel (b) reports the null rejection probabilities in simulations with $\epsilon_{i,t} \sim$ *Cauchy* and $T = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

		$p(w)$			$p(w)$		
		0.25	0.5	0.75	0.25	0.5	0.75
		(a) $\epsilon_{i,t} \sim N(0,1)$, $T = 100$			(b) $\epsilon_{i,t} \sim$ *Cauchy*, $T = 50{,}000$		
ϕ	0.25	0.052	0.047	0.054	0.031	0.031	0.034
	0.5	0.049	0.049	0.048	0.048	0.039	0.043
	0.75	0.058	0.046	0.054	0.052	0.047	0.057

Figure A13 shows the same power plots for $T = 1000$ time periods. In this case, we observe that the conservative test has good power against the weak null of no unit-i average dynamic causal effects for both $p = 0, 1$. This suggests that detecting unit-i average dynamic causal effects requires a long time dimension in the panel experiment.

*Simulation results for the estimator of the lag-*1*, unit-i weighted average dynamic causal effect,* $\bar{\tau}^{\dagger}_{i,\cdot}(1,0;1)$ We now present simulation results that analyze the properties of our

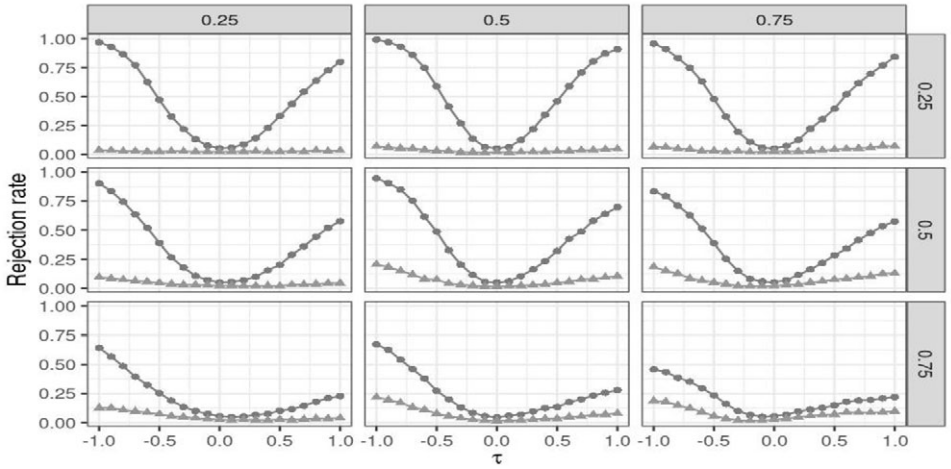


FIGURE A12. Rejection probabilities for a test of the null hypothesis $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;0) = 0$ and $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;1) = 0$ as the parameter β varies under different choices of the parameter ϕ and treatment probability $p(w)$. The rejection rate curve against $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;0) = 0$ is plotted in the circles and the rejection rate curve against $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;1) = 0$ is plotted in the triangles. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. The simulations are conducted with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 100$. Results are averaged over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

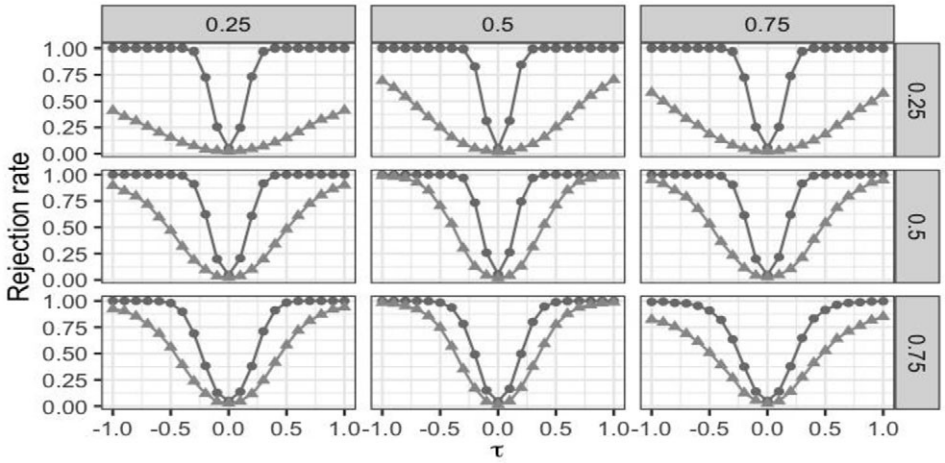
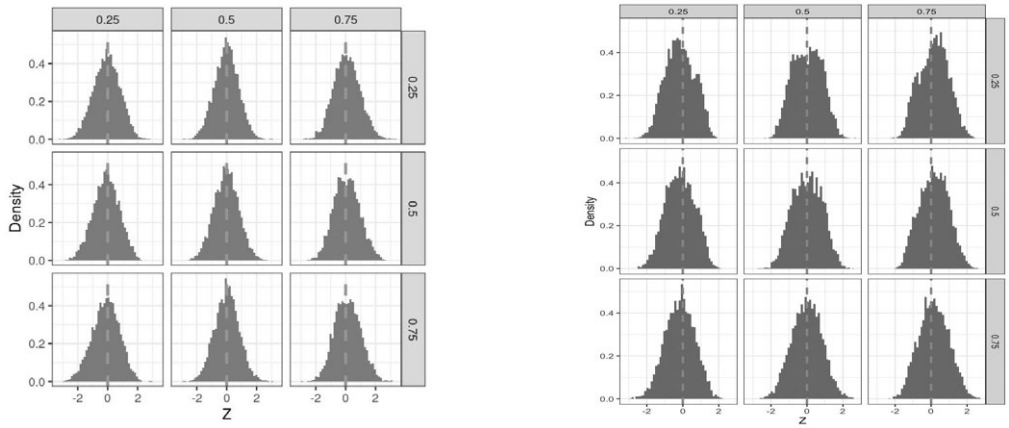
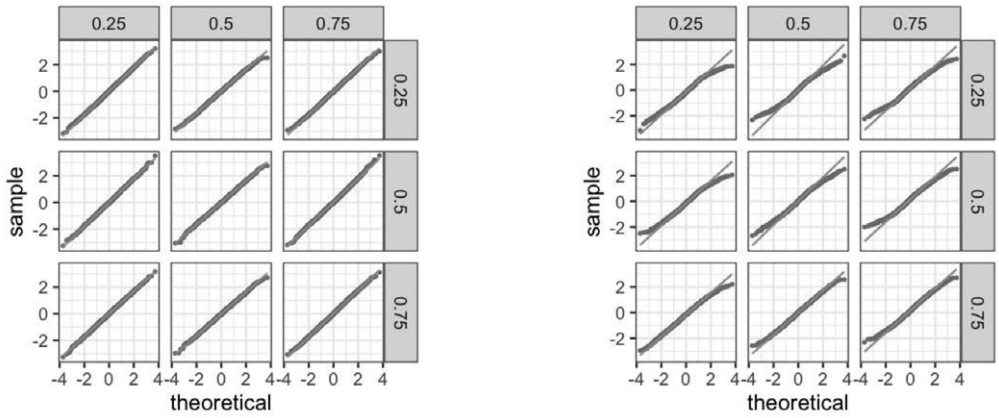


FIGURE A13. Rejection probabilities for a test of the null hypothesis $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;0) = 0$ and $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;1) = 0$ as the parameter β varies under different choices of the parameter ϕ and treatment probability $p(w)$. The rejection rate curve against $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;0) = 0$ is plotted in the circles and the rejection rate curve against $H_0 : \bar{\tau}^{\dagger}_{i\cdot}(1,0;1) = 0$ is plotted in the triangles. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. The simulations are conducted with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 1000$. Results are averaged over 5000 simulations. See Section 5 of the main text for further details on the simulation design.



(a) $\epsilon_{i,t} \sim N(0,1)$, $T = 100$ (b) $\epsilon_{i,t} \sim Cauchy$, $T = 50,000$

FIGURE A14. Simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}_{i\cdot}(1,0;1)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 100$. Panel (b) plots the simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim$ *Cauchy* and $T = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.



(a) $\epsilon_{i,t} \sim N(0,1)$, $T = 1000$ (b) $\epsilon_{i,t} \sim Cauchy$, $T = 50,000$

FIGURE A15. Quantile-quantile plots for the simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}_{i\cdot}(1,0;1)$ under different choices of the parameter ϕ and treatment probability $p(w)$. The quantile-quantile plots compare the quantiles of the simulated randomization distribution (y-axis) against the quantiles of a standard normal random variable (x-axis). The 45-degree line is plotted by the solid line. The rows index the parameter $\phi \in \{0.25, 0.5, 0.75\}$, and the columns index the treatment probability $p(w) \in \{0.25, 0.5, 0.75\}$. Panel (a) plots the quantile-quantile plots for simulated randomization distribution with normally distributed errors $\epsilon_{i,t} \sim N(0,1)$ and $T = 1000$. Panel (b) plots the quantile-quantile plots simulated randomization distribution with Cauchy distribution errors $\epsilon_{i,t} \sim Cauchy$ and $T = 50{,}000$. Results are computed over 5000 simulations. See Section 5 of the main text for further details on the simulation design.

estimator for the lag-1 total weighted average dynamic causal effect, $\hat{\bar{\tau}}^{\dagger}_{i,\cdot}(1,0;1)$. We choose the weights to $a_{\mathbf{v}}$ to place equal weight on the future treatment paths. Figure A14 plots the simulated randomization distribution for $\hat{\bar{\tau}}^{\dagger}_{i,\cdot}(1,0;1)$ and Figure A15 plots the associated quantile-quantile plot. We observe that the normal approximation remains accurate for lagged dynamic causal effects.

APPENDIX D: ADDITIONAL EMPIRICAL RESULTS

D.1 *Analysis of unit and time-specific average dynamic causal effects*

We estimate unit-specific average dynamic causal effects in the panel experiment conducted by Andreoni and Samuelson (2006). We focus on two randomly selected units in the experiment and construct estimates of their average i, tth lag-0 dynamic causal effect, $\tau_{i,t}(1,0;0)$ (Definition 5). Figure A16 shows the nonparametric estimates $\hat{\tau}_{i,t}(1,0;0)$ for $t \in [T]$, for the two units. The figure also contains the nonparametric estimate of the average unit-i lag-0 dynamic causal effect $\bar{\tau}_{i\cdot}(1,0;0) = \frac{1}{T}\sum_{t=1}^{T} \hat{\tau}_{i,t}(1,0;0)$. The result shows that the point estimate of the average unit-i lag-0 dynamic causal effect is positive for both units, suggesting that a larger value of λ in the current game increases the likelihood of cooperation for both units. Since each unit only plays a to-

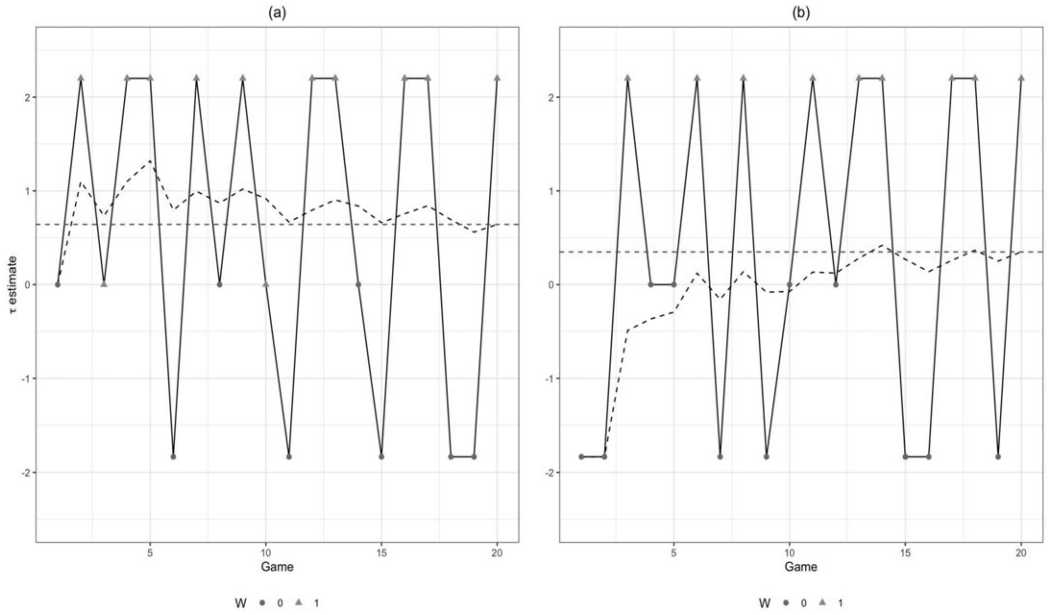


FIGURE A16. Estimates of the weighted average i, tth lag-0 dynamic causal effect (Definition 5) of $W = \mathbb{1}\{\lambda \geq 0.6\}$ on cooperation in period one for two units in the experiment of Andreoni and Samuelson (2006). The solid line plots the nonparametric estimator $\hat{\tau}_{i,t}(1,0;0)$ given in Remark 3.1. The varying dashed line plots the running average of the period-specific estimator for each unit: for each $t \in [T]$, $\frac{1}{t}\sum_{s=1}^{t} \hat{\tau}_{i,s}(1,0;0)$. The constant dashed line plots the estimated weighted average unit-i lag-0 dynamic causal effect, $\hat{\bar{\tau}}_{i\cdot}(1,0;0) = \frac{1}{T}\sum_{t=1}^{T} \hat{\tau}_{i,t}(1,0;0)$.

tal of twenty rounds, the estimated variance of these unit-specific estimators is quite large.

We next estimate period-specific, weighted average dynamic causal effects that pool information across units in order to gain precision. For each time period $t \in [T]$, we construct estimates based on the nonparametric estimator of the weighted average time-t, lag-p dynamic causal effect $\bar{\tau}^{\dagger}_{\cdot t}(1,0;p) = \frac{1}{N}\sum_{i=1}^{N} \tau^{\dagger}_{i,t}(1,0;p)$ for $p = 0,1,2,3$. For each value of p, the dashed black line in Figure A17 plots the estimates $\hat{\bar{\tau}}^{\dagger}_{\cdot t}(1,0;p)$ and the grey region plots a 95% pointwise conservative confidence band for the period-specific weighted average dynamic causal effects. For each value of p, there appears to be some heterogeneity in the period-specific weighted causal dynamic causal effects across time periods.

To further investigate these dynamic causal effects, the solid line in Figure A17 plots the nonparametric estimator the total lag-p weighted average causal effect $\bar{\tau}^{\dagger}(1,0;p)$ for $p = 0,1,2,3$, which further pools information across all units and time periods. The short dashed lines plot the conservative confidence interval for the total lag-p weighted average causal effect. See the main text for further discussion of the total lag-p weighted average causal effect estimates.

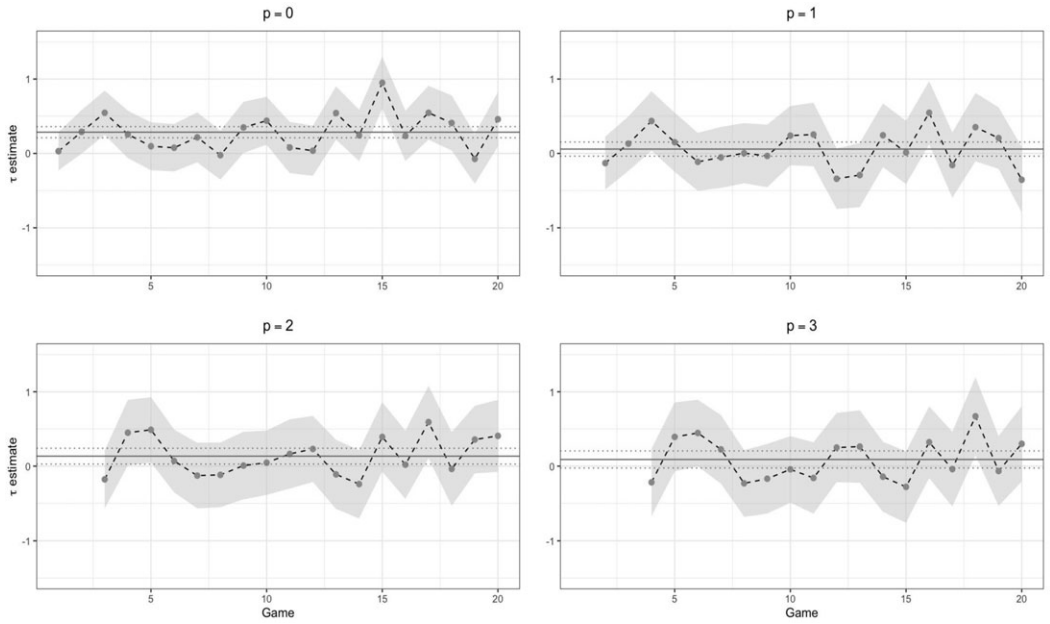


FIGURE A17. Estimates of the time-t lag-p weighted average dynamic causal effect, $\bar{\tau}^{\dagger}_{\cdot t}(1,0;\,p)$ of $W = \mathbb{1}\{\lambda \geq 0.6\}$ on cooperation in period one based on the experiment of Andreoni and Samuelson (2006) for each time period $t \in [T]$ and $p = 0, 1, 2, 3$. The long dashed line plots the nonparametric estimator of the time-t lag-p weighted average dynamic causal effect, $\hat{\bar{\tau}}^{\dagger}_{\cdot t}(1,0;\,p)$, for each period $t \in [T]$. The grey region plots the 95% pointwise confidence band for $\bar{\tau}^{\dagger}_{\cdot t}(1,0;\,p)$ based on the conservative estimator of the asymptotic variance of the nonparametric estimator (Theorem 3.2). The solid line plots the nonparametric estimator of the total lag-p weighted average dynamic causal effect, $\hat{\bar{\tau}}^{\dagger}(1,0;\,p)$ and the short dashed lines plot the 95% confidence interval for $\bar{\tau}^{\dagger}(1,0;\,p)$ based on the conservative estimator of the asymptotic variance of the nonparametric estimator.

REFERENCES

Andreoni, J. and L. Samuelson (2006), "Building rational cooperation." *Journal of Economic Theory*, 127, 117–154. [18, 19, 20]

Bojinov, I. and N. Shephard (2019), "Time series experiments and causal estimands: Exact randomization tests and trading." *Journal of the American Statistical Association*, 114 (528), 1665–1682. [3]

Hall, P. and C. C. Heyde (1980), *Martingale Limit Theory and Its Applications*. Academic Press, San Diego, CA, USA. [3]

Co-editor Andres Santos handled this manuscript.

Manuscript received 12 October, 2020; final version accepted 18 May, 2021; available online 28 May, 2021.