Semiparametric inference for impulse response functions using double/debiased machine learning

Daniele Ballinari & Alexander Wehrli

Swiss National Bank

ASSA 2026

Disclaimer: The views, opinions, findings, and conclusions or recommendations expressed in this paper are strictly those of the authors. They do not necessarily reflect the views of the Swiss National Bank (SNB). The SNB takes no responsibility for any errors or omissions in, or for the correctness of, the information contained in this paper.

Introduction

Impulse response function (IRF) for horizon $h \in \mathbb{N}_0$:

$$\theta_0^{(h)} = \mathbb{E}\left[\mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]\right]$$

where we observe $\{(Y_t, X_t, D_t) \,:\, t \in \mathcal{T}\}$ and D_t is a discrete policy variable.

Introduction

Impulse response function (IRF) for horizon $h \in \mathbb{N}_0$:

$$\theta_0^{(h)} = \mathbb{E}\left[\mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]\right]$$

where we observe $\{(Y_t, X_t, D_t) : t \in \mathcal{T}\}$ and D_t is a discrete policy variable.

- Effect of monetary policy decision on macroeconomic aggregates; $D_t \in \{-0.50, -0.25, 0.00, +0.25, +0.50\}$
- Price impact of order submission in financial markets; $D_t \in \{$"USD 1M sell", "no order submission", "USD 1M buy"$\}$

Introduction

Impulse response function (IRF) for horizon $h \in \mathbb{N}_0$:

$$\theta_0^{(h)} = \mathbb{E}\left[\mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]\right]$$

where we observe $\{(Y_t, X_t, D_t) : t \in \mathcal{T}\}$ and D_t is a discrete policy variable.

- Effect of monetary policy decision on macroeconomic aggregates; $D_t \in \{-0.50, -0.25, 0.00, +0.25, +0.50\}$
- Price impact of order submission in financial markets; $D_t \in \{$"USD 1M sell", "no order submission", "USD 1M buy"$\}$

Introduction

Impulse response function (IRF) for horizon $h \in \mathbb{N}_0$:

$$\theta_0^{(h)} = \mathbb{E}\left[\mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]\right]$$

where we observe $\{(Y_t, X_t, D_t) : t \in \mathcal{T}\}$ and D_t is a discrete policy variable.

- Effect of monetary policy decision on macroeconomic aggregates; $D_t \in \{-0.50, -0.25, 0.00, +0.25, +0.50\}$
  - *Random forest to predict inflation* (Medeiros et al., 2021)
- Price impact of order submission in financial markets; $D_t \in \{\text{``USD 1M sell''}, \text{``no order submission''}, \text{``USD 1M buy''}\}$
  - *Recurrent neural networks to predict high-frequency price changes* (Lucchese et al., 2024; Zhang et al., 2019)

Introduction

Impulse response function (IRF) for horizon $h \in \mathbb{N}_0$:

$$\theta_0^{(h)} = \mathbb{E}\left[\mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]\right]$$

where we observe $\{(Y_t, X_t, D_t) : t \in \mathcal{T}\}$ and D_t is a discrete policy variable.

- Effect of monetary policy decision on macroeconomic aggregates; $D_t \in \{-0.50, -0.25, 0.00, +0.25, +0.50\}$
  - *Random forest to predict inflation* (Medeiros et al., 2021)
- Price impact of order submission in financial markets; $D_t \in \{$"USD 1M sell", "no order submission", "USD 1M buy"$\}$
  - *Recurrent neural networks to predict high-frequency price changes* (Lucchese et al., 2024; Zhang et al., 2019)

Our contribution: bring ML-based estimation approaches for *cross-sectional* data (Chernozhukov et al., 2018) to a *time series* setting.

▸ Further references

Plug-in estimator

1. Estimate nuisance functions (conditional expectations) using an appropriate ML algorithm: $\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t]$
2. Estimate the impulse response function:

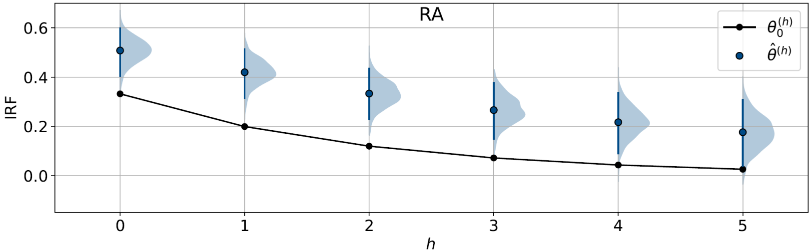
$$\hat{\theta}_{RA}^{(h)} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \hat{\mathbb{E}}[Y_{t+h}|D_t = 1, X_t] - \hat{\mathbb{E}}[Y_{t+h}|D_t = 0, X_t]$$

Plug-in estimator

1. Estimate nuisance functions (conditional expectations) using an appropriate ML algorithm: $\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t]$
2. Estimate the impulse response function:

$$\hat{\theta}^{(h)}_{RA} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \hat{\mathbb{E}}[Y_{t+h}|D_t = 1, X_t] - \hat{\mathbb{E}}[Y_{t+h}|D_t = 0, X_t]$$

Figure: True and estimated IRF across 1'000 simulated samples using Random Forest



Element 1 - Correct for plug-in bias

Recover the quantity of interest as

$$\theta_0^{(h)} = \mathbb{E}\left[g\left(Z_t^{(h)}; \Gamma_0\right)\right]$$

where $Z_t^{(h)} = (Y_{t+h}, X_t, D_t)$ and

$$g\left(Z_t^{(h)}; \Gamma_0\right) = \mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]$$

Element 1 - Correct for plug-in bias

Recover the quantity of interest as

$$\theta_0^{(h)} = \mathbb{E}\left[g\left(Z_t^{(h)};\Gamma_0\right)\right]$$

where $Z_t^{(h)} = (Y_{t+h}, X_t, D_t)$ and

$$\begin{aligned} g\left(Z_t^{(h)};\Gamma_0\right) =& \mathbb{E}[Y_{t+h}|D_t = 1, X_t] - \mathbb{E}[Y_{t+h}|D_t = 0, X_t] \\ &+ \frac{D_t}{\Pr[D_t = 1|X_t]}(Y_{t+h} - \mathbb{E}[Y_{t+h}|D_t = 1, X_t]) \\ &- \frac{1 - D_t}{1 - \Pr[D_t = 1|X_t]}(Y_{t+h} - \mathbb{E}[Y_{t+h}|D_t = 0, X_t]) \end{aligned}$$

Neyman orthogonality: small deviations from the true nuisance functions (cond. expectations and probabilities) have no first-order effect on the estimation of $\theta_0^{(h)}$ ▸ Neyman orthogonality

Neyman orthogonal estimator

1. Estimate nuisance functions using appropriate ML algorithms: $\hat{\Gamma} = \left(\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t], \widehat{\Pr}[D_t = 1|X_t]\right)$
2. Estimate the impulse response function:

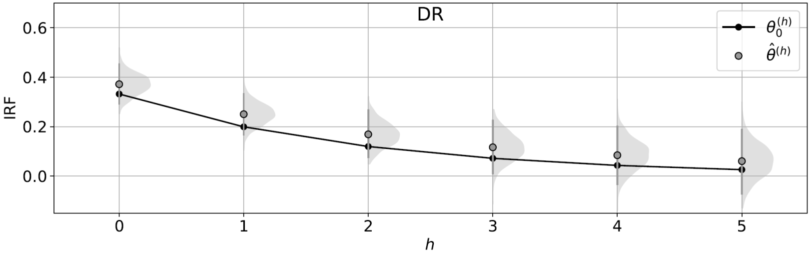
$$\hat{\theta}_{DR}^{(h)} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} g\left(Z_t^{(h)}; \hat{\Gamma}\right)$$

Neyman orthogonal estimator

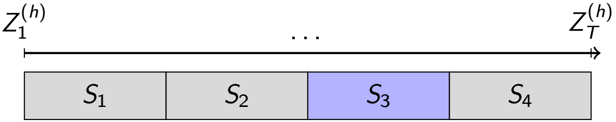
1. Estimate nuisance functions using appropriate ML algorithms:
   $\hat{\Gamma} = \left(\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t], \widehat{\Pr}[D_t = 1|X_t]\right)$
2. Estimate the impulse response function:

$$\hat{\theta}_{DR}^{(h)} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} g\left(Z_t^{(h)}; \hat{\Gamma}\right)$$

Figure: True and estimated IRF across 1'000 simulated samples using Random Forest



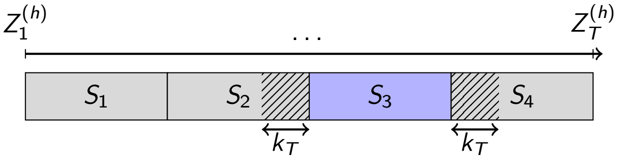
Element 2 - Cross-fitting to mitigate overfitting



1. Estimate nuisance functions using appropriate ML algorithms on S_1, S_2, S_4: $\hat{\Gamma} = \left(\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t], \widehat{\Pr}[D_t = 1|X_t]\right)$
2. Estimate the impulse response function on S_3:

$$\hat{\theta}_{S_3}^{(h)} = \frac{1}{|\mathcal{T}_3|} \sum_{t \in \mathcal{T}_3} g\left(Z_t^{(h)}; \hat{\Gamma}\right)$$

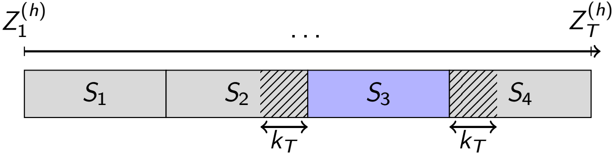
Element 2 - Cross-fitting to mitigate overfitting



1. Estimate nuisance functions using appropriate ML algorithms on S_1, S_2, S_4: $\hat{\Gamma} = \left(\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t], \widehat{\Pr}[D_t = 1|X_t]\right)$
2. Estimate the impulse response function on S_3:

$$\hat{\theta}_{S_3}^{(h)} = \frac{1}{|\mathcal{T}_3|} \sum_{t \in \mathcal{T}_3} g\left(Z_t^{(h)}; \hat{\Gamma}\right)$$

Element 2 - Cross-fitting to mitigate overfitting



1. Estimate nuisance functions using appropriate ML algorithms on S_1, S_2, S_4: $\hat{\Gamma} = \left(\hat{\mathbb{E}}[Y_{t+h}|D_t = d, X_t], \widehat{\Pr}[D_t = 1|X_t]\right)$
2. Estimate the impulse response function on S_3:

$$\hat{\theta}_{S_3}^{(h)} = \frac{1}{|\mathcal{T}_3|} \sum_{t \in \mathcal{T}_3} g\left(Z_t^{(h)}; \hat{\Gamma}\right)$$

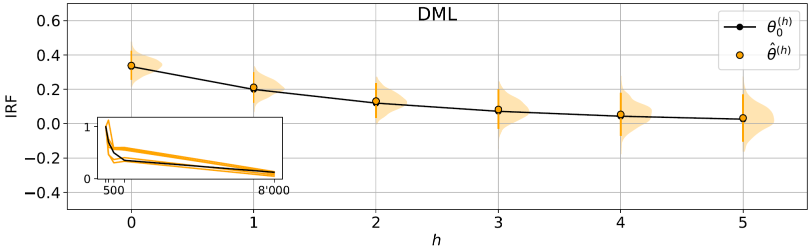
3. Repeat steps 1. and 2. and obtain IRF estimate: $\frac{1}{4}\sum_{i=1}^{4} \hat{\theta}_{S_i}^{(h)}$

Double machine learning estimator

Consistent and asymptotic normally distributed IRF estimator converging at parametric rate when combining:

- Element 1: Neyman orthogonality
- Element 2: K-fold cross-fitting

Figure: True and estimated IRF across 1'000 simulated samples using Random Forest



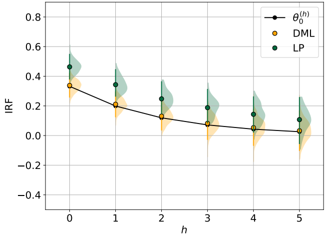
Valid inference across different sample sizes

Table: Simulation results for horizon $h = 0$ for different sample sizes

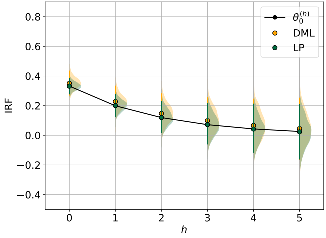
T	Neyman orthogonal + Cross-fitting				Plug-in				Neyman orthogonal			
	Bias	std($\hat{\theta}_h$)	RMSE	C_b(95%)	Bias	std($\hat{\theta}_h$)	RMSE	C_b(95%)	Bias	std($\hat{\theta}_h$)	RMSE	C_b(95%)
125	**0.053**	0.797	0.799	**0.945**	0.360	0.455	0.580	0.511	0.263	0.413	0.490	0.707
250	**0.060**	0.398	0.402	**0.949**	0.285	0.300	0.413	0.514	0.208	0.278	0.347	0.732
500	**0.030**	0.220	0.222	**0.954**	0.219	0.196	0.294	0.464	0.142	0.183	0.232	0.776
1'000	**0.029**	0.144	0.147	**0.931**	0.182	0.139	0.229	0.422	0.103	0.131	0.167	0.784
8'000	**0.006**	0.042	0.042	**0.966**	0.176	0.048	0.182	0.016	0.040	0.042	0.058	0.853

Details simulation

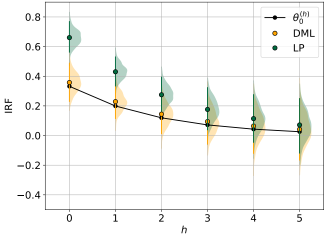
Comaprison to local projections



(a) Nonlinear DGP



(b) Linear DGP



(c) Linear DGP with interactions

Summary of key assumptions

- $g\left(Z_t^{(h)}; \Gamma_0\right)$ satisfies assumptions for the central limit theorem for α-mixing processes.
- Machine learners converge fast enough to the true nuisance functions Γ_0: product of their L_2-norm approximation error converges at rate $\sqrt{T}$.
- $Z_t^{(h)}$ is α-mixing such that after removing k_T observations between training and inference sample, they become asymptotically independent.

▸ Assumptions

Theoretical result

Theorem

Given the stochastic process $S = \{Z_t^{(h)} : t \in \mathcal{T}\}$*, define* $K \geq 2$ *sub-sequences* $S_i = \left\{Z_t^{(h)} : t \in \mathcal{T}_i\right\}$*. Define the estimator as*

$$\hat{\theta}^{(h)} = \sum_{i=1}^{K} \frac{|\mathcal{T}_i|}{T} \hat{\theta}_{S_i}^{(h)} \quad \textit{with}$$

$$\hat{\theta}_{S_i}^{(h)} = \frac{1}{|\mathcal{T}_i|} \sum_{t \in \mathcal{T}_i} g\left(Z_t^{(h)}, h; \hat{\Gamma}_{S_{-i}}\right),$$

where the nuisance functions $\hat{\Gamma}_{S_{-i}}$ *are estimated using the sub-sequence* $S_{-i} = \{Z_t^{(h)} : t \in \mathcal{T}_{-i}\}$*. It holds that*

$$\sqrt{T}(\hat{\theta}^{(h)} - \theta_0^{(h)}) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, V_0^{(h)}\right).$$

Conclusion and outlook

- Semiparametric estimator that opens up the possibility to use machine learning approaches to estimate IRFs
- Provide theoretical results of the estimator's asymptotic consistency and normality
- Simulation results and empirical application provide evidence of finite sample performance of the estimator
- Restriction: estimator requires a discrete policy variable D_t (or appropriate discretization)
- Work-in-progress: partially linear local projections estimator that relaxes this restriction

$$Y_{t+h} = \theta_0^{(h)} D_t + m(X_t) + \epsilon_{t+h}$$

References I

Adamek, R., Smeekes, S., & Wilms, I. (2024). Local Projection Inference in High Dimensions. *The Econometrics Journal*, 1–20. doi:10.1093/ectj/utae012

Angrist, J. D., Jordà, Ò., & Kuersteiner, G. M. (2018). Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited. *Journal of Business & Economic Statistics*, *36*(3), 371–387. doi:10.1080/07350015.2016.1204919

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1), C1–C68. doi:10.1111/ectj.12097

Coulombe, P. G., & Klieber, K. (2025). *Opening the black box of local projections.* (Preprint (arXiv:2505.12422))

Hauzenberger, N., Huber, F., Klieber, K., & Marcellino, M. (2025). Machine learning the macroeconomic effects of financial shocks. *Economics Letters*, *250*, 112260. doi:https://doi.org/10.1016/j.econlet.2025.112260

References II

Jordà, Ò., & Taylor, A. M. (2016). The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy. *The Economic Journal*, *126*(590), 219–255. doi:10.1111/ecoj.12332

Lucchese, L., Pakkanen, M. S., & Veraart, A. E. (2024). The short-term predictability of returns in order book markets: A deep learning perspective. *International Journal of Forecasting*, *40*(4), 1587-1621. doi:https://doi.org/10.1016/j.ijforecast.2024.02.001

Medeiros, M. C., Vasconcelos, G. F. R., Álvaro Veiga, & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, *39*(1), 98–119. doi:10.1080/07350015.2019.1637745

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708. doi:10.2307/1913610

References III

Paranhos, L. (2025). How do firms' financial conditions influence the transmission of monetary policy? a non-parametric local projection approach. *Journal of Econometrics*, *249*, 105886. doi:https://doi.org/10.1016/j.jeconom.2024.105886

Zhang, Z., Zohren, S., & Roberts, S. (2019). Deeplob: Deep convolutional neural networks for limit order books. *IEEE Transactions on Signal Processing*, *67*(11), 3001–3012.

Appendix

Further references

- Adamek, Smeekes, and Wilms (2024): inference for high-dimensional penalized local projections
- Paranhos (2025): generalized random forest for locally linear local projections in a panel setting
- Hauzenberger, Huber, Klieber, and Marcellino (2025): impulse response functions computed from Bayesian neural networks
- Angrist, Jordà, and Kuersteiner (2018) and Jordà and Taylor (2016): use inverse probability weighting to estimate IRFs
- Coulombe and Klieber (2025): plug-in estimator with random forests

Back to Introduction

Plug-in bias

Illustrative decomposition of the plug-in estimator:

$$
\begin{aligned}
\sqrt{T}\left(\hat{\theta}-\theta_0\right) = & \frac{1}{\sqrt{T}}\sum_{t\in\mathcal{T}}\left(g(Z_t;\Gamma_0)-\theta_0\right) \\
& -\frac{1}{\sqrt{T}}\sum_{t\in\mathcal{T}}\Bigg(\frac{D_t}{\widehat{\Pr}[D_t=1|X_t]}(Y_t-\hat{\mathbb{E}}[Y_t|D_t=1,X_t]) \\
& \qquad -\frac{1-D_t}{1-\widehat{\Pr}[D_t=1|X_t]}(Y_t-\hat{\mathbb{E}}[Y_t|D_t=0,X_t])\Bigg) \\
& + \text{“empirical process”} + \text{“reminder”}
\end{aligned}
$$

Neyman orthogonality: $\partial_r \mathbb{E}[g(Z_t;\Gamma_0+r(\hat{\Gamma}-\Gamma_0))]\Big|_{r=0}=0$

Back

Simulation study data generating process

$$Y_t = b(X_t) + (D_t - 0.5)\,\tau(X_t) + \gamma Y_{t-1} + \epsilon_t,$$

with

$$\begin{aligned}
e_0(X_t) &= \left(1 + e^{-X_{1,t}} + e^{-X_{2,t}}\right)^{-1} \\
b(X_t) &= 0.5\left((X_{1,t} + X_{2,t} + X_{3,t})^{+} + (X_{4,t} + X_{5,t})^{+}\right) \\
\tau(X_t) &= (X_{1,t} + X_{2,t} + X_{3,t})^{+} - (X_{4,t} + X_{5,t})^{+}
\end{aligned}$$

and

$$\begin{aligned}
\epsilon_t &= \sigma_t \zeta_t, \qquad \zeta_t \sim \mathcal{N}(0,1) \quad \text{and} \quad \sigma_t^2 = \omega + \beta_1 \zeta_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \\
X_t &= \sum_{i=1}^{p} A_i X_{t-i} + \sum_{j=1}^{q} M_j u_{t-j} + u_t, \qquad u_t \sim \mathcal{N}(0, \sigma_u^2 I_n) \\
D_t | X_t &\sim \mathrm{Ber}\left(e_0(X_t)\right)
\end{aligned}$$

Back

Assumptions I

Assumption

For some $\beta > 2$, the following conditions hold.

1. *The stochastic process $\mathcal{G}$ is weakly stationary.*
2. *For the variance it holds that* $0 < V_0^{(h)} = \lim_{T\to\infty} \mathrm{Var}\left[\frac{1}{\sqrt{T}} \sum_{t\in\mathcal{T}} g\left(Z_t^{(h)}, h; \Gamma_0\right)\right]$.
3. *$\mathcal{G}$ is uniformly L_β-bounded, i.e.* $\sup_{t\in\mathcal{T}} \mathbb{E}\left[\left|g\left(Z_t^{(h)}, h; \Gamma_0\right)\right|^\beta\right] < \infty$.
4. *$\mathcal{G}$ is α-mixing, with coefficients $\alpha(s)$, $s \in \mathbb{N}$, satisfying $\sum_{s=1}^{\infty} \alpha(s)^{(\beta-2)/\beta} < \infty$.*

Assumptions II

Assumption

Let the realization set be $\Xi_T^{(h)}$, which is a shrinking neighborhood of the true nuisance functions $\Gamma_0 = (\mu_0(d, x, h), e_0(x))$. Let $\{\Delta_T\}_{T\geq 1}$ and $\{\delta_T\}_{T\geq 1}$ be sequences of positive constants converging to zero.
Define the statistical rates $r_{\mu,T} = \sup_{t\in\mathcal{T}} \sup_{\mu\in\Xi_T^{(h)}} \|\mu(D_t, X_t, h) - \mu_0(D_t, X_t, h)\|_2$ and $r_{e,T} = \sup_{t\in\mathcal{T}} \sup_{e\in\Xi_T^{(h)}} \|e(X_t) - e_0(X_t)\|_2$. Let C be a fixed strictly positive constant. For all $i = 1, ..., K$ and $h \in \mathbb{N}_0$, the following conditions hold.

1. *The nuisance function estimators $\hat{\Gamma}_{S_{-i}}$ belong to $\Xi_T^{(h)}$ with probability at least $1 - \Delta_T$.*
2. *For $q > 2$, we have $\sup_{t\in\mathcal{T}} \sup_{\mu\in\Xi_T^{(h)}} \|\mu(D_t, X_t, h) - \mu_0(D_t, X_t, h)\|_q \leq C < \infty$ and $\sup_{t\in\mathcal{T}} \sup_{e\in\Xi_T^{(h)}} \|e(X_t) - e_0(X_t)\|_q \leq C < \infty$.*
3. $r_{\mu,T} \leq \delta_T$, $r_{e,T} \leq \delta_T$ *and* $r_{\mu,T} \cdot r_{e,T} \leq T^{-1/2}\delta_T$.
4. $\sup_{t\in\mathcal{T}} \sup_{e\in\Xi_T^{(h)}} \|e(X_t) - 1/2\|_\infty \leq 1/2 - \eta$ *for* $0 < \eta < 1$.
5. $\sup_{t\in\mathcal{T}} \mathbb{E}\left[(Y_{t+h} - \mu_0(d, X_t, h))^2 \,|X_t, D_t = d\right] \leq \epsilon_d^2 < \infty$

Assumptions III

Assumption

For all $t \in \mathcal{T}$ *and* $h \in \mathbb{N}_0$ *we have that*
$\mathbb{E}[Y_{t+h}|X_t, D_t = d, \{Z_u^{(0)} : u \in \mathcal{T}, u < t\}] = \mathbb{E}[Y_{t+h}|X_t, D_t = d]$ *and*
$\mathbb{E}[D_t|X_t, \{Z_u^{(0)} : u \in \mathcal{T}, u < t\}] = \mathbb{E}[D_t|X_t]$.

Assumptions IV

Assumption

For $d \in \{0, 1\}$, $h \in \mathbb{N}_0$ and some scalar constant $p \geq 1$ the following conditions hold.

1. $k_T = O(T)$.
2. *The nuisance functions* $\mu_0(d, x, h)$, $e_0(x)$ *and the functions* $\mu(d, x, h), e(x) \in \Xi_T^{(h)}$ *are measurable.*
3. *For for* $r > p$ *and* $1/r = 1/r' + 1/r''$, *we have* $\sup_{t \in \mathcal{T}} \|\sup_{\mu \in \Xi_T^{(h)}} (\mu(d, X_t, h) - \mu_0(d, X_t, h))\|_{2r'} < \infty$ *and* $\sup_{t \in \mathcal{T}} \|e_0(X_t) - D_t\|_{2r''} < \infty$.
4. *For* $q > p$ *and* $1/q = 1/q' + 1/q''$, *we have* $\sup_{t \in \mathcal{T}} \|\sup_{e \in \Xi_T^{(h)}} (e(X_t) - e_0(X_t))\|_{2q'} < \infty$ *and* $\sup_{t \in \mathcal{T}} \|Y_t - \mu_0(D_t, X_t, h)\|_{2q''} < \infty$.
5. *The stochastic process* $S = \{Z_t^{(h)} : t \in \mathcal{T}\}$ *is* α*-mixing with coefficients* $\alpha(s)$, *satisfying for* $T \to \infty$ *that* $\alpha(k_T)^{\psi} = o(T^{-1})$, *where* $\psi = 1/p - 1/\min(r, q)$.

Back

Variance estimator I

Assumption

The following conditions hold.

1. *There are some fixed finite constants C and $r > 4$ such that* $\sup_{t \in \mathcal{T}} \mathbb{E}\left[\left|g\left(Z_t^{(h)}, h; \Gamma_0\right)\right|^r\right] < C$.
2. *There exists a measurable function $m(z)$ such that* $\sup_{\Gamma \in \Xi_T} |g(z, h; \Gamma)| < m(z)$, *where for some finite constant D,* $\sup_{t \in \mathcal{T}} \mathbb{E}[m(Z_t)^2] < D$.
3. *For $q > 2$ and some fixed strictly positive and finite constant C we have* $\sup_{t \in \mathcal{T}} \|Y_t\|_q \leq C$.
4. *For some scalar $0 < b_m < b_r \leq 1/2$ it holds that:*
   4.1 *The bandwidth m_T is a function of the sample size such that* $\lim_{T \to \infty} m_T = \infty$ *and for $T \to \infty$ it holds that $T^{-b_m} m_T = o(1)$.*
   4.2 $r_{\mu,T} \leq \delta_T T^{-b_r}$ *and* $r_{e,T} \leq \delta_T T^{-b_r}$.

Variance estimator II

Theorem

Given the stochastic process $S = \{Z_t^{(h)} : t \in \mathcal{T}\}$, define the sub-sequences S_i for $i = 1, ..., K \geq 2$. Furthermore, define $v_t^{(h)} = g\left(Z_t^{(h)}, h; \Gamma_0\right) - \theta_0^{(h)}$. Let $w(s, m_T) = 1 - s/(m_T + 1)$, where m_T is a bandwidth parameter. Moreover, define the following Newey and West (1987) type variance estimators for each sub-sequence as

$$\widehat{V}_{S_i}^{(h)} = \frac{1}{|\mathcal{T}_i|}\left(\sum_{t\in\mathcal{T}_i}(\hat{v}_{S_{-i},t}^{(h)})^2 + 2\sum_{s=1}^{m_T} w(s, m_T)\sum_{t\in\mathcal{T}_{i,s}} \hat{v}_{S_{-i},t}^{(h)}\hat{v}_{S_{-i},t-s}^{(h)}\right).$$

The variance estimator is finally defined as

$$\widehat{V}^{(h)} = \sum_{i=1}^{K}\frac{|\mathcal{T}_i|}{T}\widehat{V}_{S_i}^{(h)}$$

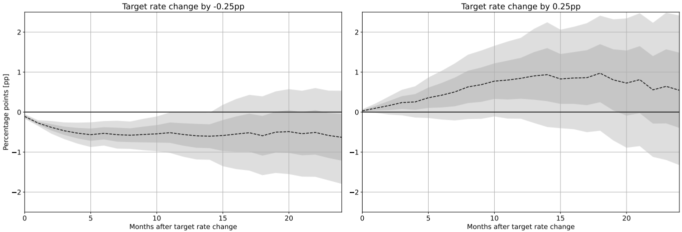
As $T \to \infty$ it holds that

$$\left|\widehat{V}^{(h)} - V_0^{(h)}\right| \xrightarrow{p} 0$$

with measure $\mathcal{P}$.

Empirical results (Angrist et al., 2018)

(a) Federal Funds Rate



(b) Unemployment Rate