Local Projection Inference is Simpler and More Robust Than You Think

José Luis Montiel Olea        Mikkel Plagborg-Møller
Columbia University            Princeton University

Gary Chamberlain Seminar
April 24, 2020
Inference on impulse responses

- **Impulse responses** are central objects for causal/counterfactual analysis in macro.
  - Response of $y_{i,t}$ due to exogenous policy change, at horizons $h = 1, 2, \ldots$

- How to do frequentist inference?

1. **Vector Autoregression (VAR):** Iterate on
   \[
   y_t = \sum_{\ell=1}^{p} A_\ell y_{t-\ell} + u_t.
   \]
   Standard errors: delta method or bootstrap.

2. **Local Projection (LP):** Jordà (2005)
   \[
   y_{i,t+h} = \beta_i(h)' y_t + \text{controls} + \xi_{i,t}(h), \quad h = 1, 2, \ldots
   \]
   Standard errors: HAC/HAR, since $\xi_{i,t}(h)$ is serially correlated.
• **Common issues in applied work:**
  1. Persistent data.
  2. Interest in long impulse response horizons $h$.

• **VAR inference:** Standard procedures break down in parts of parameter space.
  - (Near-)unit roots. Phillips (1998); Inoue & Kilian (2020)
  - Asymptotics: fixed $h$. Non-normal limit when $h = h_T \propto \sqrt{T}$ or $\propto T$. Wright (2000); Pesavento & Rossi (2006, 2007); Mikusheva (2012)

• **LP inference:** Despite widespread use, no theoretical comparisons with VARs.
  - Jordà (2005); Kilian & Kim (2011); Brugnolini (2018)
Example: Ramey (2016) handbook chapter

Gertler-Karadi monetary shock, 90\% CI. Sample: 1990m1–2012m6.

Largest horizon $h = 18\%$ of sample size $T$. 
Our contributions

- Assume VAR($p$) model.

1. Lag-augmented LP inference:

$$y_{i,t+h} = \hat{\beta}_{i}(h)'y_{t} + \sum_{\ell=1}^{p} \hat{\gamma}_{i,\ell}(h)'y_{t-\ell} + \hat{\xi}_{i,t}(h), \quad h = 1, 2, \ldots$$

2. Lag-augmented LP inference is uniformly valid over both...

   i. DGP. Includes unit root.

   ii. Horizon $h$. Includes $h = h_{T} \propto T^{\eta}$ for $\eta \in [0, 1)$ (and $\propto T$ if no unit root).

3. Lag augmentation obviates need for HAC/HAR s.e.

   - Heteroskedasticity-robust (Eicker-Huber-White) s.e. suffice.

   - Simple. No need to choose HAR procedure, tuning parameters.
Related literature

- **LP vs. VAR: identification, estimation.** Plagborg-Møller & Wolf (2019)

- **VAR inference.**
  - Lag augmentation: Toda & Yamamoto (1995); Dolado & Lütkepohl (1996); Inoue & Kilian (2020)
  - Uniformity: Mikusheva (2007, 2012); I&K (2020)

- **LP inference.**
  - Pointwise asymptotics: Jordà (2005); Kilian & Lütkepohl (2017); Stock & Watson (2018)
  - Lag augmentation: Dufour, Pelletier & Renault (2006); Breitung & Brüggemann (2020)

**This paper:** uniform LP inference, long+short horizons, simple s.e.
Outline

1. Overview of results: AR(1) case
2. Comparison with alternative methods: theory and simulations
3. Formal uniformity result: AR(1) case
4. General VAR($p$) case
5. Empirical illustration
6. Conclusion
Model

- Start with univariate AR(1) model for clarity:

\[ y_t = \rho y_{t-1} + u_t, \quad t = 1, 2, \ldots, T, \quad y_0 = 0. \]

- Parameter of interest: impulse response at horizon \( h \).

\[ \beta(\rho, h) \equiv \rho^h, \quad \rho \in [-1, 1], \quad h \in \mathbb{N}. \]

**Assumption 1: Mean independence**

\( \{u_t\} \) is strictly stationary, and \( E(u_t \mid \{u_s\}_{s \neq t}) = 0. \)

- Stronger than MDS: \( E(u_t \mid \{u_s\}_{s < t}) = 0. \)

- Satisfied for i.i.d. \( u_t \). Also allows many types of heteroskedasticity/SV.
Non-augmented local projection: fragile, HAR s.e.

- **AR(1) model implies**
  \[ y_{t+h} = \beta(\rho, h) y_t + \xi_t(\rho, h) \]
  \[ \equiv \rho^h \quad \equiv \sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell} \]

- **Non-augmented LP estimator**: regress \( y_{t+h} \) on \( y_t \) (no controls).
  - Consistent and asy. normal when \( |\rho| \ll 1 \).
  - Non-normal limit when \( \rho \approx 1 \) since \( y_t \) non-stationary.
  - Requires HAR s.e. even when \( |\rho| \ll 1 \), since \( \xi_t(\rho, h) \) serially correlated. HAR inference challenging in small samples. Involves tuning parameters. Müller (2007, 2014); Lazarus, Lewis, Stock & Watson (2018)
Lag-augmented local projection: robust inference

• Lag-augmented LP:

\[
\begin{pmatrix}
\hat{\beta}(h) \\
\hat{\gamma}(h)
\end{pmatrix}
\equiv
\left(\sum_{t=1}^{T-h} x_t x_t'\right)^{-1} \sum_{t=1}^{T-h} x_t y_{t+h}, \quad x_t \equiv (y_t, y_{t-1})'.
\]

• Would get same \(\hat{\beta}(h)\) if we regressed on \((u_t, y_{t-1})\), since \(u_t = y_t - \rho y_{t-1}\).

• \(\hat{\beta}(h)\) has uniform normal limit, since

\[
y_{t+h} = \beta(\rho, h) u_t + \beta(\rho, h + 1) y_{t-1} + \xi_t(\rho, h).
\]

• \(\hat{\gamma}(h)\) non-normal when \(\rho \approx 1\), but we don’t care.
Lag-augmented local projection: simple standard errors

\[ y_{t+h} = \beta(\rho, h)u_t + \beta(\rho, h+1)y_{t-1} + \xi_t(\rho, h) \equiv \sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell} \]

- **Bonus:** lag augmentation simplifies standard errors.

- **Leading term in asymptotic expansion:**

  \[ \hat{\beta}(h) \approx \beta(\rho, h) + \frac{\sum_{t=1}^{T-h} \xi_t(\rho, h)u_t}{\sum_{t=1}^{T-h} u_t^2}. \]

- \( \xi_t(\rho, h) \) is serially correlated, but scores \( \xi_t(\rho, h)u_t \) are not: For \( s < t \),

  \[ E[\xi_t(\rho, h)u_t \xi_s(\rho, h)u_s] = E[\xi_t(\rho, h)u_t \xi_s(\rho, h) E(u_s \mid u_{s+1}, u_{s+2}, \ldots)] = 0 \]

- Requires \( E(u_t \mid \{u_s\}_{s>t}) = 0 \). MDS is not enough.
Lag-augmented local projection: robust inference

- Heteroskedasticity-robust (Eicker-Huber-White) s.e. $\hat{s}(h)$ suffice. No tuning param’s.

- Define usual confidence interval:

$$\hat{C}(h, \alpha) \equiv \left[ \hat{\beta}(h) - z_{1-\alpha/2} \hat{s}(h), \hat{\beta}(h) + z_{1-\alpha/2} \hat{s}(h) \right].$$

- Proposition: This CI is uniformly valid.

$$\inf_{\rho \in [-1,1]} \inf_{1 \leq h \leq \bar{h}_T} P_{\rho} \left( \beta(\rho, h) \in \hat{C}(h, \alpha) \right) \to 1 - \alpha,$$

for any seq \( \{\bar{h}_T\} \in \mathbb{N} \) such that $\bar{h}_T / T \to 0$.

- Further result: If we restrict $|\rho| \leq 1 - a$ for $a > 0$, then even $\bar{h}_T \propto T$ is OK.

- Non-normal limit for $\rho = 1$, $h = h_T \propto T$. 
Outline

1. Overview of results: AR(1) case
2. Comparison with alternative methods: theory and simulations
3. Formal uniformity result: AR(1) case
4. General VAR($p$) case
5. Empirical illustration
6. Conclusion
Fragility of AR inference

• Simple lag-aug LP inference robust to persistence $\rho$ and horizon $h$. Not true for textbook AR delta method inference. Phillips (1998); Benkwitz et al. (2000); Pesavento & Rossi (2007)

• Lag-augmented AR: $\hat{\beta}_{\text{ARLA}}(h) \equiv \hat{\rho}_1^h$, where $y_t = \hat{\rho}_1 y_{t-1} + \hat{\rho}_2 y_{t-2} + \hat{u}_t$.
  
  • Uniformly $\sqrt{T}$-normal limit for fixed $h$.
  
  • Efron bootstrap CI valid at long horizons. Inoue & Kilian (2020)

  • But estimator is inconsistent at horzs $h = h_T \geq \kappa \sqrt{T}$ when $\rho \approx 1$. Confidence interval does not shrink with $T$ (length can even explode), unlike LP.

• AR grid bootstrap valid at short+long horizons, but not intermediate. Computationally intensive. Hansen (1999); Mikusheva (2012)
Relative efficiency in stationary, fixed-horizon case

- Consider a stationary, homoskedastic VAR model. Fix the horizon $h$.
- Then textbook non-augmented AR estimator achieves semiparametric efficiency bound.
- **Ambiguous** ranking of asymptotic variances of inefficient procedures:
  1. Lag-augmented AR.
  2. Lag-augmented LP.
  3. Non-augmented LP (requires HAC standard errors).
- In paper: Ranking in homoskedastic AR(1) model, as function of $(\rho, h)$. 

Simulation study

• Confidence interval procedures:
  1. Non-augmented AR, delta method s.e. (straw man).
  2. Lag-augmented AR, Efron bootstrap CI. Inoue & Kilian (2020)
  3. Non-augmented LP, percentile-t bootstrap CI, HAR s.e.
  4. Lag-augmented LP, percentile-t bootstrap CI, EHW s.e.

• Bootstrap: wild recursive AR design.

• AR(1) model. \( T = 240 \). Nominal confidence level: 90%.

• \( u_t \sim N(0, 1) \) i.i.d. (ARCH innovations qualitatively similar.)
<table>
<thead>
<tr>
<th>$h$</th>
<th>Coverage</th>
<th>Median length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$AR_d$</td>
<td>$AR^*_b$</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.897</td>
<td>0.899</td>
</tr>
<tr>
<td>6</td>
<td>0.829</td>
<td>0.899</td>
</tr>
<tr>
<td>12</td>
<td>0.758</td>
<td>0.899</td>
</tr>
<tr>
<td>36</td>
<td>0.628</td>
<td>0.899</td>
</tr>
<tr>
<td>60</td>
<td>0.581</td>
<td>0.899</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.848</td>
<td>0.902</td>
</tr>
<tr>
<td>6</td>
<td>0.808</td>
<td>0.902</td>
</tr>
<tr>
<td>12</td>
<td>0.765</td>
<td>0.902</td>
</tr>
<tr>
<td>36</td>
<td>0.655</td>
<td>0.902</td>
</tr>
<tr>
<td>60</td>
<td>0.585</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.532</td>
<td>0.882</td>
</tr>
<tr>
<td>6</td>
<td>0.494</td>
<td>0.882</td>
</tr>
<tr>
<td>12</td>
<td>0.454</td>
<td>0.882</td>
</tr>
<tr>
<td>36</td>
<td>0.348</td>
<td>0.882</td>
</tr>
<tr>
<td>60</td>
<td>0.288</td>
<td>0.882</td>
</tr>
</tbody>
</table>

$\rho = 0.50$

$\rho = 0.95$

$\rho = 1.00$
Outline

1. Overview of results: AR(1) case
2. Comparison with alternative methods: theory and simulations
3. Formal uniformity result: AR(1) case
4. General VAR(\(p\)) case
5. Empirical illustration
6. Conclusion
Additional regularity assumptions

Assumption 2: Moments

\( E(u_t^8) < \infty \), and \( E(u_t^2 \mid \{u_s\}_{s<t}) \geq \delta > 0 \).

\( \{u_t^2\} \) has absolutely summable cumulants up to 4th order.

Assumption 3: Contiguity

For any sequence \( \{\rho_T\} \in [-1, 1] \) such that \( \lim_{T \to \infty} T(1 - \rho_T) < \infty \), the sequence of probability measures \( \{P_{\rho_T}\} \) is contiguous to \( P_1 \).

- Used to derive OLS convergence rates when \( \rho \approx 1 \). Requires smoothness of conditional density. Jeganathan (1995); Jansson (2008)
Proposition 1: Uniform inference

Let Ass’t “Mean independence” & “Moments” hold. Let \( a \in (0, 1) \).

i) For all \( x \in \mathbb{R} \),

\[
\sup_{\rho \in [-1+a, 1-a]} \sup_{1 \leq h \leq (1-a)T} \left| P_{\rho} \left( \frac{\hat{\beta}(h) - \beta(\rho, h)}{\hat{s}(h)} \leq x \right) - \Phi(x) \right| \rightarrow 0.
\]

ii) Let additionally Ass’n “Contiguity” hold. Consider any sequence \( \{\tilde{h}_T\} \in \mathbb{N} \) such that \( \tilde{h}_T / T \rightarrow 0 \). Then for all \( x \in \mathbb{R} \),

\[
\sup_{\rho \in [-1+a, 1]} \sup_{1 \leq h \leq \tilde{h}_T} \left| P_{\rho} \left( \frac{\hat{\beta}(h) - \beta(\rho, h)}{\hat{s}(h)} \leq x \right) - \Phi(x) \right| \rightarrow 0.
\]

Remark: Non-normal limit when \( \rho = 1 \) and \( h = h_T \propto T \) (like AR).
Uniformly valid LP inference: key proof challenges

- Study all drifting sequences \{\rho_T, h_T\}. Andrews, Cheng & Guggenberger (2019)

  \[
  \hat{\beta}(h_T) - \beta(\rho_T, h_T) \propto \sum_{t=1}^{T-h_T} \xi_t(\rho_T, h_T)u_t + \sum_{t=1}^{T-h_T} \xi_t(\rho_T, h_T) \left[ \hat{u}_t(h_T) - u_t \right] = [\rho_T - \hat{\rho}(h_T)] y_{t-1}
  \]

- 1st term: \( E[\xi_t(\rho_T, h_T)^2 u_t^2] \rightarrow \infty \) if \( \rho_T \rightarrow 1, h_T \rightarrow \infty \).
  - CLT for MDS. Must “reverse time” b/c \( E[\xi_t(\rho, h)u_t | \{u_s\}_{s<t}] \neq 0 \).
  - Explicitly calculate uniform moment bounds.

- 2nd term: Convergence rate of \( \hat{\rho}(h_T) \) depends on whether \( \rho_T \approx 1 \). Mikusheva (2007)
  - Explicit moment calculations when \( \rho_T \ll 1 \).
  - When \( \rho_T \approx 1 - c/T \), appeal to contiguity as’n. Need then only consider familiar case \( \rho = 1 \).
Outline

1. Overview of results: AR(1) case
2. Comparison with alternative methods: theory and simulations
3. Formal uniformity result: AR(1) case
4. General VAR($p$) case
5. Empirical illustration
6. Conclusion
VAR(p) model

- VAR(p) model for $n$-dimensional data vector:
  \[
  y_t = \sum_{\ell=1}^{p} A_{\ell} y_{t-\ell} + u_t, \quad t = 1, \ldots, T, \quad y_0 = \cdots = y_{1-p} = 0.
  \]

- Still impose conditional mean independence: $E(u_t \mid \{u_s\}_{s \neq t}) = 0$.

- Known lag length $p$.

- Reduced-form impulse responses of $y_{1,t}$ at horizon $h$: $\beta_1(A, h) \in \mathbb{R}^n$.

- Parameter of interest: $\nu' \beta_1(A, h)$, where $\nu \in \mathbb{R}^n$, $\nu \neq 0$.
  - Simple extension: joint inference on vector $\beta_1(A, h)$.
  - Extensions for future work: structural impulse responses, deterministic dynamics.
Multivariate lag-augmented local projection

- VAR model implies: Jordà (2005); Kilian & Lütkepohl (2017)

\[ y_{1,t+h} = \beta_1(A, h)' y_t + \sum_{\ell=1}^{p-1} \delta_{1,\ell}(A, h)' y_{t-j} + \xi_{1,t}(A, h), \quad \xi_{1,t}(A, h) \equiv \sum_{\ell=1}^{h} \beta_1(A, h - \ell)' u_{t+\ell}. \]

- Lag-augmented LP estimator: regression with \( p \) lags as controls.

\[ y_{1,t+h} = \hat{\beta}_1(h)' y_t + \sum_{\ell=1}^{p} \hat{\delta}_{1,\ell}(h)' y_{t-j} + \hat{\xi}_{1,t}(A, h). \]

- Confidence interval for \( \nu' \beta_1(A, h) \):

\[ \hat{C}_1(h, \nu, \alpha) \equiv \nu' \hat{\beta}_1(h) \pm z_{1-\alpha/2} \hat{s}_1(h, \nu). \]
Uniformly valid multivariate LP inference

- **Proposition:** Impose As’n “Mean independence”. Then the CI \( \hat{C}_1(h, \nu, \alpha) \) is asymptotically uniformly valid over the DGP and horizon \( h \), provided...

1. The VAR(\( p \)) model can be written in the form \( y_t = \Lambda y_{t-1} + \tilde{y}_t \), where: Phillips (1988)
   - \( \tilde{y}_t \) is uniformly stationary VAR(\( p - 1 \)) (geometrically decaying IRFs).
   - \( \Lambda = \text{diag}(\rho_1, \ldots, \rho_n) \) with \( \rho_i \in [-1, 1] \). Mikusheva (2012)

2. Either:
   - i) \( h \leq (1 - a) T \) and \( |\rho_i| \leq 1 - a \), \( i = 1, \ldots, n \), where \( a > 0 \). OR:
   - ii) \( h \leq \bar{h}_T \), where \( \bar{h}_T / T \rightarrow 0 \).

3. Further regularity conditions on \{u_t\} hold.
Uniformly valid multivariate LP inference: discussion

- We are not aware of other uniformity results that allow multiple (near-)unit roots.
- VAR proof follows AR(1) intuition. Main challenge: uniform bounds on IRFs.
- **Corollary**: Can allow for cointegration among control variables $y_{2,t}, \ldots, y_{n,t}$.
- Possibly non-normal limit when $\rho_i \approx 1$ for some $i$ and $h \propto T$. 


Outline

1. Overview of results: AR(1) case
2. Comparison with alternative methods: theory and simulations
3. Formal uniformity result: AR(1) case
4. General VAR($p$) case
5. Empirical illustration
6. Conclusion
Response of Excess Bond Premium to 25 bp monetary shock

Outline

1. Overview of results: AR(1) case
2. Comparison with alternative methods: theory and simulations
3. Formal uniformity result: AR(1) case
4. General VAR($p$) case
5. Empirical illustration
6. Conclusion
Conclusion


• We show that lag-augmented LP inference on impulse responses is robust to:
  1. Persistence of data.
  2. Length of impulse response horizon.

• Efficiency loss for stationary DGPs at short horizons, but modest in absolute terms.

• Lag augmentation obviates need for HAR s.e. Simple!

• Only known VAR-based methods with comparable robustness are either computationally demanding or can yield very long CIs. Mikusheva (2012); Inoue & Kilian (2020)
Conclusion


- We show that lag-augmented LP inference on impulse responses is robust to:
  1. Persistence of data.
  2. Length of impulse response horizon.

- Efficiency loss for stationary DGPs at short horizons, but modest in absolute terms.

- Lag augmentation obviates need for HAR s.e. Simple!

- Only known VAR-based methods with comparable robustness are either computationally demanding or can yield very long CIs. Mikusheva (2012); Inoue & Kilian (2020)

Thank you!
Heteroskedastic innovations

- Assume $u_t = \tau_t \epsilon_t$, where $\tau_t \geq 0$.
- Assume $\epsilon_t$ is i.i.d., $E(\epsilon_t) = 0$.
- Then $E(u_t \mid \{u_s\}_{s \neq t}) = 0$ if either...
  - $\{\tau_t\} \perp \perp \{\epsilon_t\}$ (SV).
  - $\tau_t = f(\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots)$ and distribution of $\epsilon_t$ is symmetric (GARCH).
Lag-augmented local projection: simple standard errors

- Heteroskedasticity-robust (Eicker-Huber-White) s.e. suffice:

\[
\hat{s}(h) \equiv \frac{(\sum_{t=1}^{T-h} \hat{\xi}_t(h)^2 \hat{u}_t(h)^2)^{1/2}}{\sum_{t=1}^{T-h} \hat{u}_t(h)^2},
\]

where

\[
\hat{\xi}_t(h) \equiv y_{t+h} - \hat{\beta}(h)y_t - \hat{\gamma}(h)y_{t-1},
\]

\[
\hat{u}_t(h) \equiv y_t - \hat{\rho}(h)y_{t-1},
\]

\[
\hat{\rho}(h) \equiv (\sum_{t=1}^{T-h} y_t y_{t-1})/(\sum_{t=1}^{T-h} y_{t-1}^2).
\]

- Readily computed by standard statistical software.

- No tuning parameters.
AR inference: medium-long horizons

- Suppose $\rho > 0$ and we use some asymptotic normal estimator $\hat{\rho}$:

  $$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{d} N(0, \tau^2).$$

- Delta method s.e.:

  $$\text{se}(\hat{\rho}) \equiv h|\hat{\rho}^{h^{-1}\hat{\tau}}| \sqrt{T}, \quad \hat{\tau} \overset{p}{\to} \tau.$$

- Then at horizon $h = h_T = \sqrt{T}$,

  $$\frac{\hat{\rho}^{h_T}}{\rho^{h_T}} = e^{\sqrt{T}(\log \hat{\rho} - \log \rho)} \xrightarrow{d} e^{N(0, \tau^2/\rho^2)},$$

  $$\left| \frac{\sqrt{T}(\hat{\rho}^{h_T} - \rho^{h_T})}{h_T \hat{\rho}^{h_T-1\hat{\tau}}} \right| = \left| \hat{\rho}^{h_T} \left( 1 - \frac{\hat{\rho}^{h_T}}{\rho^{h_T}} \right) \right| \xrightarrow{d} \frac{\rho}{\tau} \left( 1 - N(0, \tau^2/\rho^2) \right).$$
Trade-off $LP_{LA}$ vs. $AR_{LA}$:

- Non-linear transformation $\rho^h$.

Trade-off $LP_{LA}$ vs. $LP_{NA}$:

- Effective regressor $u_t$ vs. $y_t$.
- Serial correl’n of $\xi_t(\rho, h)y_t$. 
Proof sketch: reversing time

\[ \xi_t(\rho, h) \equiv \sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell} \]

- Run sum “backwards in time”:
  \[ \sum_{t=1}^{T-h_T} \xi_t(\rho_T, h_T) u_t = \sum_{t=1}^{T-h_T} \chi_{T,t}, \quad \chi_{T,t} \equiv \xi_{T-h-T+1}(\rho_T, h_T)u_{T-h-T+1}. \]

- Define filtration

  \[ \mathcal{F}_{T,t} \equiv \sigma(u_{T-h_T-t+1}, u_{T-h_T-t+2}, \ldots). \]

  Then \( \chi_{T,t} \in \mathcal{F}_{T,t} \) and \( \mathcal{F}_{T,t} \subset \mathcal{F}_{T,t+1} \) for all \( t \).

- \( \{\chi_{T,t}, \mathcal{F}_{T,t}\} \) is a martingale difference array:

  \[
  E(\chi_{T,t} \mid \mathcal{F}_{T,t-1}) = \xi_{T-h_T-t+1}(\rho_T, h_T) E(u_{T-h_T-t+1} \mid \{u_{T-h_T-t+s}\}_{s>1}) = 0 \text{ by As'n "Mean independence"}
  \]
Simulation results: delta method procedures

<table>
<thead>
<tr>
<th>$h$</th>
<th>$\text{Coverage}$</th>
<th>$\text{Median length}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{AR}_d$</td>
<td>$\text{AR}^*_d$</td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.50$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.897</td>
<td>0.902</td>
</tr>
<tr>
<td>6</td>
<td>0.829</td>
<td>0.857</td>
</tr>
<tr>
<td>12</td>
<td>0.758</td>
<td>0.792</td>
</tr>
<tr>
<td>36</td>
<td>0.628</td>
<td>0.667</td>
</tr>
<tr>
<td>60</td>
<td>0.581</td>
<td>0.625</td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.95$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.848</td>
<td>0.899</td>
</tr>
<tr>
<td>6</td>
<td>0.808</td>
<td>0.880</td>
</tr>
<tr>
<td>12</td>
<td>0.765</td>
<td>0.841</td>
</tr>
<tr>
<td>36</td>
<td>0.655</td>
<td>0.732</td>
</tr>
<tr>
<td>60</td>
<td>0.585</td>
<td>0.684</td>
</tr>
<tr>
<td></td>
<td>$\rho = 1.00$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.532</td>
<td>0.890</td>
</tr>
<tr>
<td>6</td>
<td>0.494</td>
<td>0.854</td>
</tr>
<tr>
<td>12</td>
<td>0.454</td>
<td>0.813</td>
</tr>
<tr>
<td>36</td>
<td>0.348</td>
<td>0.693</td>
</tr>
<tr>
<td>60</td>
<td>0.288</td>
<td>0.633</td>
</tr>
</tbody>
</table>
Simulation results: ARCH innovations ($\alpha_1 = 0.7$)

<table>
<thead>
<tr>
<th>$h$</th>
<th>Coverage</th>
<th>Median length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$AR_d$</td>
<td>$AR^*_b$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\rho = 0.50$

|     |        |        |        |        |        |        |        |        |
| 1   | 0.870  | 0.833 | 0.907 | 0.887 | 0.292  | 0.335 | 0.332 | 0.385 |
| 6   | 0.758  | 0.833 | 0.906 | 0.905 | 0.046  | 0.085 | 0.273 | 0.245 |
| 12  | 0.678  | 0.833 | 0.904 | 0.902 | 0.001  | 0.007 | 0.264 | 0.239 |
| 36  | 0.554  | 0.833 | 0.896 | 0.898 | 0.000  | 0.000 | 0.279 | 0.255 |
| 60  | 0.520  | 0.833 | 0.901 | 0.908 | 0.000  | 0.000 | 0.301 | 0.274 |

$\rho = 0.95$

|     |        |        |        |        |        |        |        |        |
| 1   | 0.848  | 0.830 | 0.818 | 0.894 | 0.087  | 0.338 | 0.084 | 0.396 |
| 6   | 0.797  | 0.830 | 0.835 | 0.895 | 0.354  | 1.744 | 0.383 | 0.627 |
| 12  | 0.750  | 0.830 | 0.836 | 0.872 | 0.468  | 3.915 | 0.603 | 0.724 |
| 36  | 0.632  | 0.830 | 0.856 | 0.867 | 0.291  | 64.157 | 0.806 | 0.716 |
| 60  | 0.571  | 0.830 | 0.887 | 0.892 | 0.095  | 1028.366 | 0.880 | 0.708 |

$\rho = 1.00$

|     |        |        |        |        |        |        |        |        |
| 1   | 0.566  | 0.838 | 0.839 | 0.894 | 0.041  | 0.331 | 0.040 | 0.385 |
| 6   | 0.521  | 0.838 | 0.858 | 0.879 | 0.224  | 2.089 | 0.240 | 0.689 |
| 12  | 0.472  | 0.838 | 0.842 | 0.853 | 0.391  | 5.683 | 0.469 | 0.914 |
| 36  | 0.363  | 0.838 | 0.756 | 0.740 | 0.676  | 194.037 | 1.175 | 1.386 |
| 60  | 0.304  | 0.838 | 0.711 | 0.653 | 0.733  | 6503.801 | 1.670 | 1.496 |
VAR parameter space

- Let there be given constants $a \in [0, 1)$, $C > 0$, and $\epsilon \in (0, 1)$.

- $A(a, C, \epsilon) \equiv$ space of autoregressive coefficients $A = (A_1, \ldots, A_p)$ such that the associated lag polynomial $A(L) = I_n - \sum_{\ell=1}^p A_\ell L^\ell$ admits the factorization

  $$A(L) = B(L)(I_n - \text{diag}(\rho_1, \ldots, \rho_n)L).$$

- $\rho_i \in [a - 1, 1 - a]$ for all $i = 1, \ldots, n$.

- $B(L)$ is a lag polynomial of order $p - 1$ with companion matrix $\bf{B}$ satisfying $\|B^\ell\| \leq C(1 - \epsilon)^\ell$ for all $\ell = 1, 2, \ldots$. 