Local Projections and VARs Estimate the Same Impulse Responses

Mikkel Plagborg-Møller
Princeton University

Christian K. Wolf
Princeton University

NBER Summer Institute

July 9, 2019

Estimation of IRFs

- How to estimate **impulse response functions (IRFs)**?

$$E(y_{t+h} \mid \varepsilon_{j,t} = 1) - E(y_{t+h} \mid \varepsilon_{j,t} = 0), \quad h = 0, 1, 2, \ldots$$

- Two popular competing semi-structural approaches:

1. **Structural Vector Autoregression (SVAR)**: Sims (1980)

$$A(L)w_t = B\varepsilon_t, \quad A(L) = I_n - \sum_{\ell=1}^{p} A_\ell L^\ell, \quad \varepsilon_t \sim WN(0, I_n).$$

2. **Local Projections (LP)**: Jordà (2005)

$$y_{t+h} = \mu_h + \beta_h \varepsilon_{j,t} + \text{controls} + \xi_{h,t}, \quad h = 0, 1, 2, \ldots$$

SVAR vs. LP: State of the literature

- Conventional wisdom:
  - SVAR is "more efficient". LP is "more robust to misspecification".
  - LP requires that we observe a measure of the "shock". SVAR needed for more exotic identification approaches (long-run/sign restrictions, etc.).
- Simulation studies offer conflicting rankings. Meier (2005); Kilian & Kim (2011); Brugnolini (2018); Nakamura & Steinsson (2018); Choi & Chudik (2019)
- SVAR and LP approaches often yield different empirical conclusions. Ramey (2016)

Our contributions

1. Proposition: In population, linear LPs and SVARs estimate the *same* IRFs.
   - Nonparametric result. Only requires unrestricted lag structures.
     Jordà (2005); Kilian & Lütkepohl (2017)
2. Derive implications for...
   - Efficient estimation.
   - Structural identification.
   - Identification using IV/proxy.
   - Linear estimands in nonlinear DGPs.

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing "SVAR" identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

Equivalence result: Nonparametric assumptions

- Observed data: $w_t = (\overbrace{r_t}^{n_r \times 1}{}', \overbrace{x_t}^{1\times 1}, \overbrace{y_t}^{1\times 1}, \overbrace{q_t}^{n_q \times 1}{}')'$.

- Interested in response of y_t to an impulse in x_t. Other var's: "controls" (more soon).

Assumption: Nonparametric regularity

$\{w_t\}$ is covariance stationary and purely non-deterministic, with an everywhere nonsingular spectral density matrix and absolutely summable Wold coefficients.

To simplify notation, we proceed as if $\{w_t\}$ were a (strictly stationary) jointly Gaussian vector time series.

- No assumption (yet) about underlying causal structure.

- Gaussianity: use conditional expectation/variance. Can replace with projections.

Equivalence result: Definition of LP IRF

- Consider for each $h = 0, 1, 2, \dots$ the linear projection

$$y_{t+h} = \mu_h + \beta_h x_t + \gamma_h' r_t + \sum_{\ell=1}^{\infty} \delta_{h,\ell}' w_{t-\ell} + \xi_{h,t}.$$

- $\xi_{h,t}$: projection residual.
- $\mu_h, \beta_h, \gamma_h, \delta_{h,1}, \delta_{h,2}, \dots$: projection coefficients.
- **LP IRF** of y_t with respect to x_t: $\{\beta_h\}_{h \geq 0}$.
- Note: Projection controls for the contemporaneous value of r_t but not of q_t. Also controls for all lags of all series.

Equivalence result: Definition of SVAR IRF

- Consider the multivariate linear "VAR(∞)" projection

$$w_t = c + \textstyle\sum_{\ell=1}^{\infty} A_\ell w_{t-\ell} + u_t.$$

- $u_t \equiv w_t - E(w_t \mid \{w_\tau\}_{-\infty<\tau<t})$: projection residual. $c, A_1, A_2, \dots$: proj. coefficients.
- Cholesky decomp.: $\Sigma_u \equiv E(u_t u_t') = BB'$, where B lower triangular.
- Corresponding recursive SVAR representation w. orthogonal "shocks":

$$A(L)w_t = c + B\eta_t, \quad A(L) \equiv I - \textstyle\sum_{\ell=1}^{\infty} A_\ell L^\ell, \quad \eta_t \equiv B^{-1}u_t.$$

  Note: r_t ordered first, q_t ordered last.

- **VAR IRF** of y_t with respect to an innovation in x_t: $\{\theta_h\}_{h\geq 0}$, where

$$\theta_h = C_{h,n_r+2,\bullet} B_{\bullet,n_r+1}, \quad \textstyle\sum_{\ell=0}^{\infty} C_\ell L^\ell = C(L) \equiv A(L)^{-1}.$$

Equivalence result

Proposition: Equivalence between LP and SVAR

Under Assumption "Nonparametric Regularity", the LP and VAR IRFs are equal, up to a constant of proportionality:

$$\theta_h = \sqrt{E(\tilde{x}_t^2)} \times \beta_h \quad \text{for all } h = 0, 1, 2, \dots,$$

where

$$\tilde{x}_t \equiv x_t - E(x_t \mid r_t, \{w_\tau\}_{-\infty<\tau<t}).$$

- Any LP IRF can be obtained as an appropriately ordered SVAR IRF. Ordering corresponds to contemporaneous control variables in LP.
- Constant of proportionality does not depend on y_t or h.

Equivalence result: Intuition

- Intuition: Impulse responses are just linear projections. ▸ Proof
  i) VAR impulse response: h-step least-squares forecast based on model-implied second moments.
  ii) VAR(∞) captures all second moments of data.
  $\Rightarrow$ VAR(∞) impulse response: direct projection (LP).
- Extension in paper: non-recursive SVARs.
  - Arbitrary SVAR IRF $=$ LP on a linear combination $b'w_t$ (and lags). ▸

Equivalence result: Finite lag length

- Let $\theta_h(p)$ and $\beta_h(p)$ denote the VAR and LP impulse response estimand at horizon h when we project on only p lags of the data w_t.

Proposition: Equivalence between LP and SVAR, finite lag length

Let "Nonparametric Regularity" assumption hold. Define

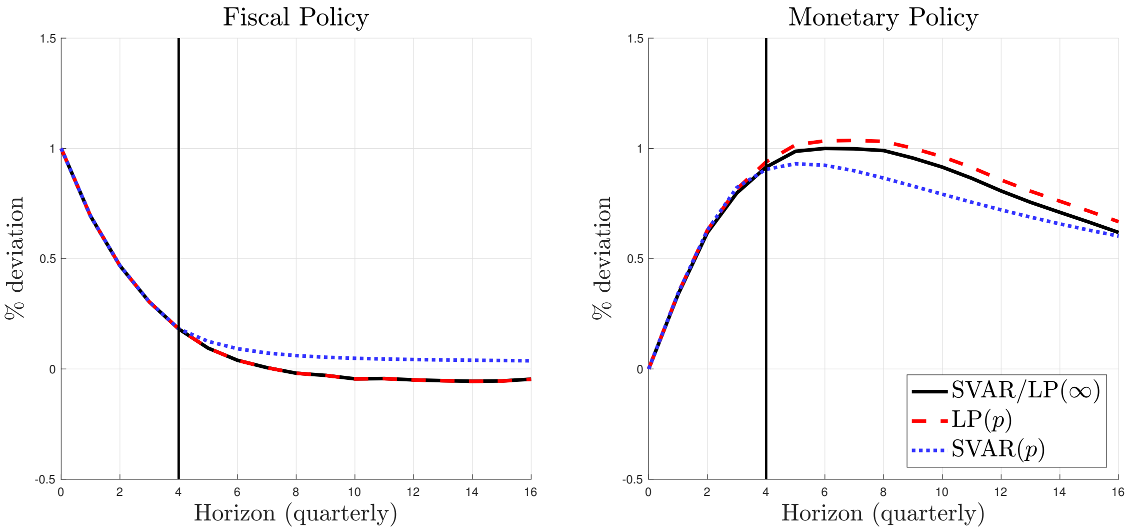
$$\tilde{x}_t(\ell) \equiv x_t - E(x_t \mid r_t, \{w_\tau\}_{t-\ell \leq \tau < t}), \quad \ell = 0, 1, 2, \ldots.$$

Let the nonnegative integers h, p satisfy $h \leq p$.

If $\tilde{x}_t(p) = \tilde{x}_t(p-h)$, then $\theta_h(p) = \sqrt{E(\tilde{x}_t(p)^2)} \times \beta_h(p)$.

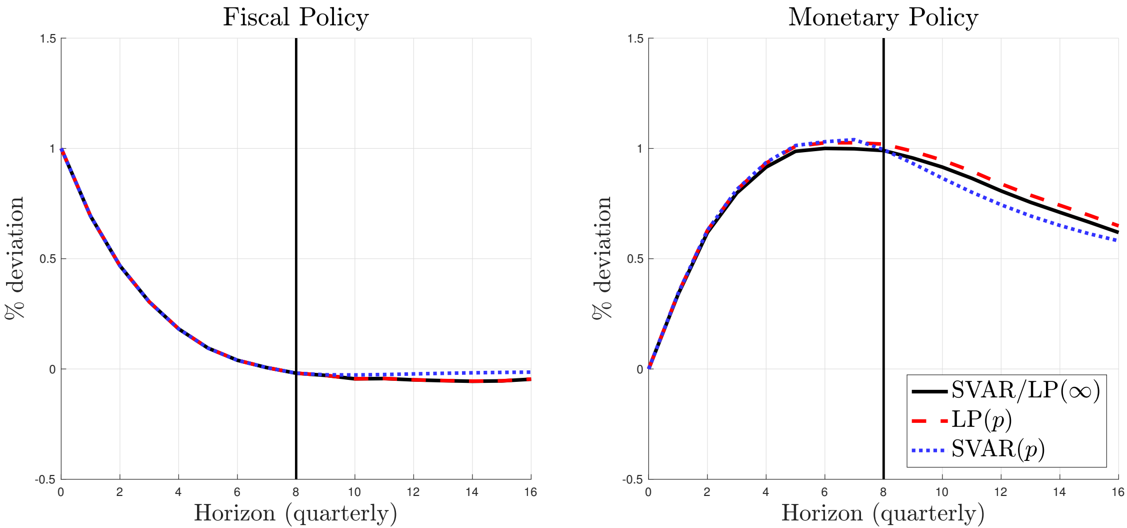
- If x_t is a "shock" that doesn't affect r_t on impact, then $\tilde{x}_t(\ell) = x_t$ for all $\ell \geq 0$.
- More generally, in practice, we often have $\tilde{x}_t(p) \approx \tilde{x}_t(p-h)$ for $h \ll p$.

Illustration: IRFs of output in Smets-Wouters model



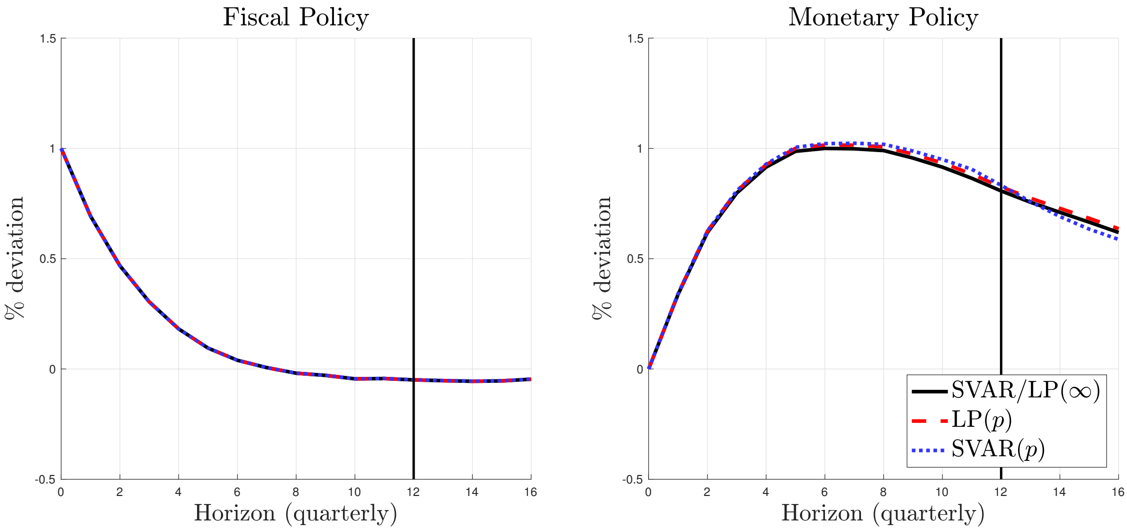
Note: $p =$ 4. Left panel: shock observed. Right panel: recursive ID.

Illustration: IRFs of output in Smets-Wouters model



Note: $p =$ 8. Left panel: shock observed. Right panel: recursive ID.

Illustration: IRFs of output in Smets-Wouters model



Note: $p = 12$. Left panel: shock observed. Right panel: recursive ID.

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing "SVAR" identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

Efficient estimation: Bias/variance trade-off

- Proposition (in paper): Sample LP and VAR estimators equivalent as $p, T \to \infty$.
- We face finite-sample bias-variance trade-off. Which dimension reduction technique for linear projection works best?
  - If DGP = VAR(p), SVAR estimator has small bias and extrapolates efficiently. Unrealistic.
  - Forecasting literature has already examined the bias-variance trade-off of "LP" vs. "VAR": direct vs. iterated multi-step forecasts.
    Schorfheide (2005); Marcellino, Stock & Watson (2006); Chevillon (2007); McElroy (2015)
  - There exists spectrum of "shrinkage" techniques: Bayes, model averaging, smoothness priors.
    Giannone, Lenza & Primiceri (2015); Hansen (2016); Plagborg-Møller (2016); Barnichon & Brownlees (2018); Miranda-Agrippino & Ricco (2018)
  - No method uniformly dominates in terms of MSE. Depends on DGP.

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing "SVAR" identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing "SVAR" identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

Structural identification: SVAR vs. LP

- Assume causal model: **Structural Vector Moving Average**. Stock & Watson (2018)

$$w_t = \mu + \textstyle\sum_{\ell=0}^{\infty} \Theta_\ell \varepsilon_{t-\ell},$$
$$\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n_\varepsilon,t})' \overset{i.i.d.}{\sim} N(0, I_{n_\varepsilon}).$$

- For now, assume all shocks are **invertible** (SVAR assumption):

$$\varepsilon_{j,t} \in \overline{\text{span}}\left(\{w_\tau\}_{-\infty<\tau\leq t}\right), \quad j = 1, 2, \dots, n_\varepsilon.$$

- Main result $\Longrightarrow$ Recursive SVAR identification can be implemented through LPs.

- Other "SVAR" identification schemes also implementable using LPs. Next: long-run ID.

Structural identification: Long-run restrictions

- Data: $w_t \equiv (\Delta gdp_t, unr_t)'$, log GDP growth and unemployment rate.
- Assume SVMA model with $n_\varepsilon = 2$ shocks. $\varepsilon_{1,t}$: supply shock, $\varepsilon_{2,t}$: demand shock.
- Assume $\sum_{\ell=0}^{\infty} \Theta_{1,2,\ell} = 0$. No long-run effect of demand shock on the *level* of output.
  Blanchard & Quah (1989)
- Given a large horizon H, consider the linear projection

$$gdp_{t+H} - gdp_{t-1} = \tilde{\mu}_H + \sum_{\ell=0}^{\infty} \tilde{\delta}'_{H,\ell} w_{t-\ell} + \tilde{\xi}_{H,t}.$$

- Proposition: $\Theta_{i,1,h} \propto \lim_{H\to\infty} \bar{\beta}_{h,H}$ for $h \geq 0$, where

$$w_{i,t+h} = \bar{\mu}_{h,H} + \bar{\beta}_{h,H}(\tilde{\delta}_{H,0}{}' w_t) + \sum_{\ell=1}^{\infty} \bar{\delta}'_{h,H,\ell} w_{t-\ell} + \bar{\xi}_{h,H,t}.$$

Implementing "SVAR" identification using LP: Summary

- SVAR identification approaches work if and only if corresponding LP approaches work.
  - Additional example in paper: sign restrictions.
- Lesson: Choice of identification approach is logically+practically distinct from choice of dimension reduction technique (i.e., linear projection estimator).
- Finite-sample bias/variance trade-off depends on specifics of DGP.

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing "SVAR" identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

LP-IV

- Popular applied strategy: Identify IRFs using **proxy/IV** z_t for $\varepsilon_{1,t}$:

$$z_t = c_z + \sum_{\ell=1}^{\infty}(\Psi_\ell z_{t-\ell} + \Lambda_\ell w_{t-\ell}) + \alpha\varepsilon_{1,t} + v_t,$$

where $v_t \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$ and independent of ε_t at all leads/lags.

LP-IV

- Popular applied strategy: Identify IRFs using **proxy/IV** z_t for $\varepsilon_{1,t}$:

$$z_t = c_z + \sum_{\ell=1}^{\infty}(\Psi_\ell z_{t-\ell} + \Lambda_\ell w_{t-\ell}) + \alpha\varepsilon_{1,t} + v_t,$$

where $v_t \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$ and independent of ε_t at all leads/lags.

- **LP-IV**: Given SVMA+IV, can estimate relative structural IRF using 2SLS version of LP:

$$y_{t+h} = \mu_h + \beta_h x_t + \sum_{\ell=1}^{\infty} \delta'_{h,\ell} w_{t-\ell} + \xi_{h,t} \quad \text{with } z_t \text{ as IV for } x_t.$$

Mertens (2015); Jordà et al. (2015, 2018); Ramey & Zubairy (2018); Stock & Watson (2018)

- Reason: $\text{Cov}(y_{t+h}, z_t \mid \{w_\tau, z_\tau\}_{-\infty<\tau<t}) = \alpha \times \Theta_{n_r+2,1,h} \quad \Longrightarrow \quad \frac{\Theta_{n_r+2,1,h}}{\Theta_{n_r+1,1,0}}$ identified.

LP-IV

- Popular applied strategy: Identify IRFs using **proxy/IV** z_t for $\varepsilon_{1,t}$:

$$z_t = c_z + \sum_{\ell=1}^{\infty}(\Psi_\ell z_{t-\ell} + \Lambda_\ell w_{t-\ell}) + \alpha\varepsilon_{1,t} + v_t,$$

where $v_t \overset{i.i.d.}{\sim} N(0, \sigma_v^2)$ and independent of ε_t at all leads/lags.

- **LP-IV**: Given SVMA+IV, can estimate relative structural IRF using 2SLS version of LP:

$$y_{t+h} = \mu_h + \beta_h x_t + \sum_{\ell=1}^{\infty} \delta'_{h,\ell} w_{t-\ell} + \xi_{h,t} \quad \text{with } z_t \text{ as IV for } x_t.$$

Mertens (2015); Jordà et al. (2015, 2018); Ramey & Zubairy (2018); Stock & Watson (2018)

- Reason: $\text{Cov}(y_{t+h}, z_t \mid \{w_\tau, z_\tau\}_{-\infty<\tau<t}) = \alpha \times \Theta_{n_r+2,1,h} \quad \Longrightarrow \quad \frac{\Theta_{n_r+2,1,h}}{\Theta_{n_r+1,1,0}}$ identified.

- $\varepsilon_{1,t}$ allowed to be non-invertible: $\varepsilon_{1,t} \notin \overline{\text{span}}(\{w_\tau\}_{-\infty<\tau\leq t})$.

LP-IV: Estimand

- Will now reinterpret LP-IV estimand. Set $W_t \equiv (z_t, w_t')'$.
- "Reduced-form" LPs:

$$y_{t+h} = \mu_{RF,h} + \beta_{RF,h} z_t + \sum_{\ell=1}^{\infty} \delta_{RF,h,\ell}' W_{t-\ell} + \xi_{RF,h,t}, \quad h \geq 0.$$

- "First-stage" LP (doesn't depend on h):

$$x_t = \mu_{FS} + \beta_{FS} z_t + \sum_{\ell=1}^{\infty} \delta_{FS,\ell}' W_{t-\ell} + \xi_{FS,t}.$$

- As usual (one IV, one endogenous covariate), 2SLS estimand given by ratio

$$\beta_{LPIV,h} \equiv \frac{\beta_{RF,h}}{\beta_{FS}}, \quad h \geq 0.$$

- Equivalence result $\implies$ $\beta_{LPIV,h}$ can be obtained from an SVAR.

LP-IV: Equivalence with recursive SVAR

Proposition: Equivalence of LP-IV and recursive SVAR

Let "Nonparametric Regularity" assumption hold for expanded data $W_t \equiv (z_t, w_t')'$.

Consider a recursive SVAR(∞) in W_t, with z_t ordered first. Define:

- $\tilde{\theta}_{y,h}$: SVAR-implied imp. resp. of y_t wrt. first shock at horizon h.
- $\tilde{\theta}_{x,0}$: SVAR-implied imp. resp. of x_t wrt. first shock *on impact*.

Then $\beta_{LPIV,h} = \tilde{\theta}_{y,h}/\tilde{\theta}_{x,0}$.

- Under structural SVMA+IV as'ns: Consistently estimate relative IRFs by ordering IV first in recursive SVAR ("internal instrument"). Robust to non-invertibility! Noh (2018)

LP-IV: Equivalence with recursive SVAR

Proposition: Equivalence of LP-IV and recursive SVAR

Let "Nonparametric Regularity" assumption hold for expanded data $W_t \equiv (z_t, w_t')'$.

Consider a recursive SVAR(∞) in W_t, with z_t ordered first. Define:

- $\tilde{\theta}_{y,h}$: SVAR-implied imp. resp. of y_t wrt. first shock at horizon h.
- $\tilde{\theta}_{x,0}$: SVAR-implied imp. resp. of x_t wrt. first shock *on impact*.

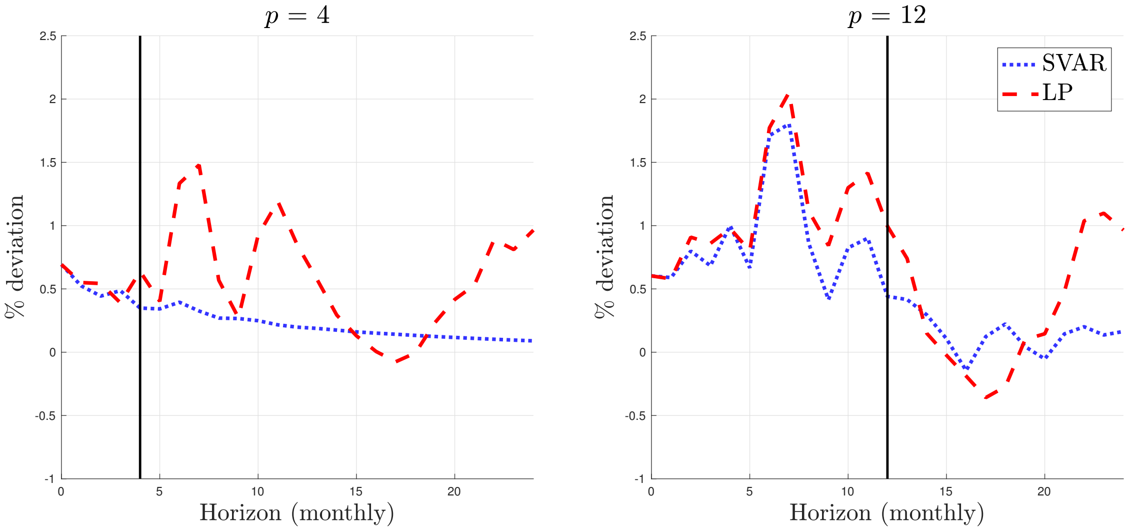
Then $\beta_{LPIV,h} = \tilde{\theta}_{y,h}/\tilde{\theta}_{x,0}$.

- Under structural SVMA+IV as'ns: Consistently estimate relative IRFs by ordering IV first in recursive SVAR ("internal instrument"). Robust to non-invertibility! Noh (2018)
- In contrast, SVAR-IV ("external instruments") estimator requires invertibility.
  Stock (2008); Stock & Watson (2012); Mertens & Ravn (2013); P-M & W (2019)

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing “SVAR” identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

Response of bond spread to monetary shock: VAR and LP estimates



Note: Shock normalized to increase 1-year bond rate by 100 basis points on impact.

Outline

1. Main equivalence result
2. Estimation
3. Structural identification
   - Implementing "SVAR" identification using LP
   - Identification with instruments
4. Empirical illustration
5. Conclusion

Conclusion

- Linear LPs and SVARs share the same population IRF estimand. Nonparametric result.
- Implications:
  - Unavoidable bias/variance trade-off in finite samples. Estimation procedures lie on a spectrum.
  - Identification $\perp$ dimension reduction. "SVAR" identification can be phrased in terms of LPs.
  - LP-IV estimator can be implemented by ordering IV/proxy first in SVAR ("internal instruments"). Robust to non-invertibility, unlike SVAR-IV ("external instruments").
  - In paper: Linear LP/VAR IRF estimand = "best linear approximation" in non-linear DGP.
- This is all about IRFs. Variance/historical decomp's more involved. P-M & W (2019)

Thank you!

LP vs. SVAR: High-freq. identification of monetary shocks

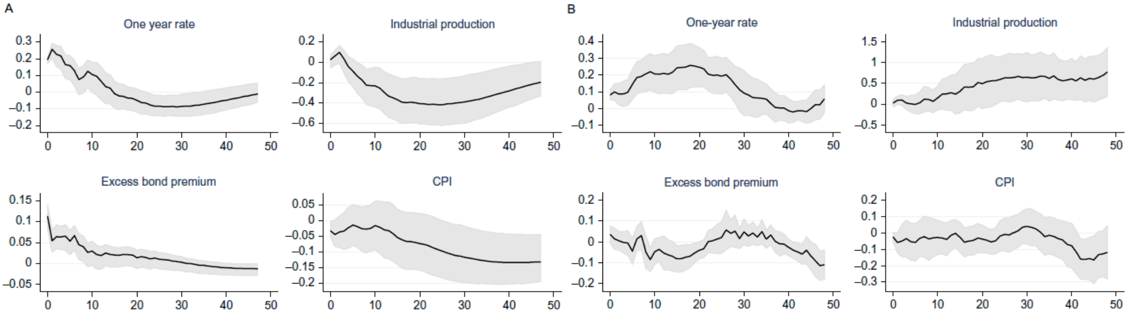


Fig. 3 Gertler–Karadi's monetary shock. (A) Gertler–Karadi's monetary proxy SVAR, VAR from 1979m7 to 2012m6, instrument from 1991m1 to 2012m6. (B) Gertler–Karadi monetary shock, Jordà 1990m1–2012m6. Light gray bands are 90% confidence bands.

Source: Ramey (2016) handbook chapter

Back

Equivalence result: Proof sketch

- Formal proof just requires least-squares algebra.
- LP estimand from Frisch-Waugh Theorem:

$$\beta_h = \frac{\text{Cov}(y_{t+h}, \tilde{x}_t)}{E(\tilde{x}_t^2)}.$$

Equivalence result: Proof sketch (cont.)

- VAR reduced-form impulse responses $A(L)^{-1}$ from Wold decomp.:

$$w_t = \chi + C(L)u_t = \chi + \textstyle\sum_{\ell=0}^{\infty} C_\ell B \eta_t, \quad \chi \equiv C(1)c.$$

- Hence, VAR estimand equals

$$\theta_h = C_{h,n_r+2,\bullet} B_{\bullet,n_r+1} = \mathrm{Cov}(y_{t+h}, \eta_{x,t}),$$

where we partition $\eta_t = (\eta'_{r,t}, \eta_{x,t}, \eta_{y,t}, \eta'_{q,t})'$.

- By $u_t = B\eta_t$ and properties of Cholesky decomposition,

$$\eta_{x,t} \propto \tilde{u}_{x,t},$$

where we partition $u_t = (u'_{r,t}, u_{x,t}, u_{y,t}, u'_{q,t})'$ and define

$$\tilde{u}_{x,t} \equiv u_{x,t} - E(u_{x,t} \mid u_{r,t}) = \tilde{x}_t.$$

- Conclude

$$\theta_h \propto \mathrm{Cov}(y_{t+h}, \tilde{x}_t) \propto \beta_h.$$

Back

Equivalence result: Non-recursive specifications

- In general, any SVAR identification scheme studies the propagation of *some* rotation of the Wold innovations:
$$\bar{\eta}_t \equiv b' u_t.$$

- Can show that the SVAR IRF to this innovation corresponds to coefficients $\{\bar{\beta}_h\}_{h\geq 0}$ from linear projections
$$y_{t+h} = \bar{\mu}_h + \bar{\beta}_h(b' w_t) + \textstyle\sum_{\ell=1}^{\infty} \bar{\delta}'_{h,\ell} w_{t-\ell} + \bar{\xi}_{h,t},$$
up to constant of proportionality.

- Equivalent LP projects on linear combination $b' w_t$ of variables.

Back

Sample asymptotic equivalence

- Consider least-squares sample analogues of LP and VAR. Include p lags of w_t in both methods.
- $\hat{x}_t(p)$: residual from regression of x_t on intercept, r_t, $w_{t-1}, \dots, w_{t-p}$.
- LP estimator (from Frisch-Waugh theorem):

$$\hat{\beta}_h(p) = \frac{\sum_{t=p+1}^{T-h} y_{t+h}\hat{x}_t(p)}{\sum_{t=p+1}^{T-h} \hat{x}_t(p)^2}.$$

- $\hat{\theta}_h(p)$: horizon-h impulse response of y_t to an innovation in x_t in a Cholesky-identified estimated VAR(p) model (with intercept).
- Will now show that $\hat{\beta}_h(p) \approx \text{constant} \times \hat{\theta}_h(p)$ as $T \to \infty$, provided $p \to \infty$ at appropriate rate.

Sample asymptotic equivalence (cont.)

Proposition: In-sample near-equivalence of LP and SVAR

Suppress notation $p = p(T)$. Assume:

i) $\{w_t\}$ is covariance stationary and has a VAR(∞) representation with $\sum_{\ell=1}^{\infty} \|A_\ell\| < \infty$. Wold innovations u_t have finite and pos. def. cov. matrix Σ. (Perhaps non-Gaussian.)

ii) Reduced-form least-squares VAR estimator satisfies

$$\|\hat{c}(p) - c\| = o_p(1), \;\; \|\hat{A}(p) - A(p)\| = o_p(1), \;\; \|\hat{\Sigma}(p) - \Sigma\| = o_p(1).$$

Lewis & Reinsel (1985); Gonçalves & Kilian (2007)

Then as $p, T \to \infty$,

$$\hat{\theta}_h(p) = \left(\tfrac{1}{T-p} \textstyle\sum_{t=p+1}^{T} \hat{x}_t(p)^2\right)^{-1/2} \times \hat{\beta}_h(p) + O_p(\hat{R}(p)),$$

$$\hat{R}(p) \equiv \frac{\max\{1, \sup_{1 \le t \le T} \|w_t\|\}^2}{T-p} + \left(\textstyle\sum_{\ell=p-h+1}^{p} \|\hat{A}_\ell(p)\|^2\right)^{1/2}.$$

Back

Structural identification: Short-run restrictions

- "Fast-r-slow" short-run identification of monetary policy shocks: CEE (2005)

$$A(L)\begin{pmatrix} r_t \\ x_t \\ q_t \end{pmatrix} = \begin{pmatrix} B_{11}\varepsilon_{1,t} \\ B_{21}\varepsilon_{1,t} + B_{22}\varepsilon_{2,t} \\ B_{31}\varepsilon_{1,t} + B_{32}\varepsilon_{2,t} + B_{33}\varepsilon_{3,t} \end{pmatrix}.$$

  ($n = 3$ for clarity.)

- x_t: Federal Funds Rate. r_t: "slow-moving". q_t: "fast-moving".

- Given this model, our equivalence result implies that the IRF of q_t (say) wrt. $\varepsilon_{2,t}$ is proportional to $\{\beta_h\}_{h\geq 0}$ from the LP

$$q_{t+h} = \mu_h + \beta_h x_t + \gamma_h r_t + \sum_{\ell=1}^{\infty} \delta'_{h,\ell} w_{t-\ell} + \xi_{h,t}.$$

Back

Long-run restrictions: Proof sketch

$$w_t = \chi + C(L)u_t, \quad u_t = B\varepsilon_t \tag{$\dagger$}$$

- Standard argument: Long-run restriction $\Theta_{1,2}(1) = 0$ implies
$$e_1' C(1)u_t = \Theta_{1,1}(1) \times \varepsilon_{1,t}.$$
- Since $w_{1,t} = \Delta gdp_t$,
$$\tilde{\delta}_H' = \text{Cov}(gdp_{t+H} - gdp_{t-1}, u_t)\Sigma_u^{-1} = \textstyle\sum_{\ell=0}^{H} \text{Cov}(w_{1,t+\ell}, u_t)\Sigma_u^{-1}.$$
- Wold decomposition ($\dagger$) implies
$$\textstyle\sum_{\ell=0}^{\infty} \text{Cov}(w_{t+\ell}, u_t)\Sigma_u^{-1} = \sum_{\ell=0}^{\infty} C_\ell = C(1).$$
- So
$$\lim_{H\to\infty} \tilde{\delta}_H' = e_1' C(1).$$
- Finally, apply main equivalence result.

Back

Structural identification: Sign restrictions

- Want IRF of y_t wrt. monetary shock. Assume SVMA + invertibility.
- Impulse response at horizon h given by $\nu' \check{\beta}_h$ for unknown $\nu \in \mathbb{R}^{n_w}$, where $\check{\beta}_h$ is obtained from projection
$$y_{t+h} = \check{\mu}_h + \check{\beta}_h' w_t + \sum_{\ell=1}^{\infty} \check{\delta}_{h,\ell}' w_{t-\ell} + \check{\xi}_{h,t}.$$
- Impose sign restrictions: r_t responds *positively* to a monetary shock at all horizons $s = 0, 1, \dots, \bar{H}$. Uhlig (2005)
- For $s = 0, 1, \dots, \bar{H}$, store coef. vector $\ddot{\beta}_s$ from projection
$$r_{t+s} = \ddot{\mu}_s + \ddot{\beta}_s' w_t + \sum_{\ell=1}^{\infty} \ddot{\delta}_{s,\ell}' w_{t-\ell} + \ddot{\xi}_{s,t}.$$
- *Largest possible* response of y_{t+h} to a monetary shock that raises r_t by one unit on impact:
$$\sup_{\nu \in \mathbb{R}^{n_w}} \nu' \check{\beta}_h \quad \text{subject to} \quad \ddot{\beta}_0' \nu = 1, \;\; \ddot{\beta}_s' \nu \geq 0, \;\; s = 1, \dots, \bar{H}.$$

Back

Examples of IVs/proxies

- Narrative monetary shocks. Romer & Romer (2004)
- Narrative fiscal shocks. Mertens & Ravn (2013); Ramey & Zubairy (2017); Mertens & M. Olea (2018)
- High-frequency asset price changes around FOMC announcements. Barakchian & Crowe (2013); Gertler & Karadi (2015)
- Oil supply disruptions. Hamilton (2003)
- Large oil discoveries. Arezki, Ramey & Sheng (2016)
- Utilization-adjusted TFP growth. Fernald (2014); Caldara & Kamps (2017)
- Volatility spikes. Carriero et al. (2015)

Back

LP-IV: Intuition for equivalence

- Why does recursive SVAR work even under non-invertibility?
- Shock $\varepsilon_{1,t}$ still non-invertible wrt. *expanded* info set:

$$\varepsilon_{1,t} \notin \text{span}\left(\{w_\tau, z_\tau\}_{-\infty<\tau\leq t}\right) \quad \text{in general.}$$

- But remaining non-invertibility is due only to classical measurement error in

$$\tilde{z}_t \equiv z_t - E\left(z_t \mid \{w_\tau, z_\tau\}_{-\infty<\tau<t}\right) = \alpha\varepsilon_{1,t} + v_t.$$

- Attenuation bias is the same (in pct terms) for all horizons and response variables $\Longrightarrow$ *Relative* impulse responses $\frac{\Theta_{n_r+2,1,h}}{\Theta_{n_r+1,1,0}}$ correctly identified (not absolute).

Back

LP-IV: Comparison with SVAR-IV

- The alternative **SVAR-IV** approach manipulates the Wold innovations $u_t \equiv w_t - E(w_t \mid \{w_\tau\}_{-\infty<\tau<t})$ from an SVAR in w_t alone.
- Specifically, SVAR-IV identifies the shock of interest as

$$\tilde{\varepsilon}_{1,t} \equiv \frac{1}{\sqrt{\text{Var}(\tilde{z}_t^\dagger)}} \times \tilde{z}_t^\dagger,$$

  where

$$\tilde{z}_t^\dagger \equiv E(\tilde{z}_t \mid u_t).$$

- $\tilde{\varepsilon}_{1,t} \neq \varepsilon_{1,t}$, except if the shock is invertible. Plagborg-Møller & Wolf (2019)

Back

Estimands in non-linear models

- Often claimed that LP is "robust to misspecification/non-linearities". Our equivalence result implies that this is not true.

- Assume the general non-linear DGP (assumed stationary)

$$w_t = g(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots), \quad \varepsilon_t \overset{i.i.d.}{\sim} N(0, I_{n_\varepsilon}).$$

- Using Wold decomposition, can represent as *linear* SVMA model

$$w_t = \mu^* + \textstyle\sum_{\ell=0}^{\infty} \Theta_\ell^* \varepsilon_{t-\ell} + \sum_{\ell=0}^{\infty} \Psi_\ell^* \zeta_{t-\ell}.$$

- ζ_t: n_w-dimensional white noise, uncorrelated at all leads/lags with ε_t.

- Linear SVMA impulse responses Θ_ℓ^* corresponding to the structural shocks ε_t have a **best linear approximation** interpretation:

$$(\Theta_0^*, \Theta_1^*, \dots) \in \underset{(\tilde{\Theta}_0, \tilde{\Theta}_1, \dots)}{\operatorname{argmin}} \; E\left[\left(g(\varepsilon_t, \varepsilon_{t-1}, \dots) - \textstyle\sum_{\ell=0}^{\infty} \tilde{\Theta}_\ell \varepsilon_{t-\ell}\right)^2\right].$$

Estimands in non-linear models (cont.)

$$(\Theta_0^*, \Theta_1^*, \dots) \in \underset{(\tilde{\Theta}_0, \tilde{\Theta}_1, \dots)}{\operatorname{argmin}} \; E\left[\left(g(\varepsilon_t, \varepsilon_{t-1}, \dots) - \textstyle\sum_{\ell=0}^{\infty} \tilde{\Theta}_\ell \varepsilon_{t-\ell}\right)^2\right]$$

- Linear SVAR/LP IRF estimand can be given "best linear approximation" interpretation.
- Estimators that rely on higher moments are not as easy to interpret under misspecification.
- We do not take a stand on whether the best linear approximation is structurally interesting. Depends on application.
- In some applications, non-linearities may be the key objects of interest, in which case *linear* SVAR/LP methods are not useful.

Back