Local Asymptotic Minimax Inference for Set-Identified Impulse Responses from Local Projection Instrumental Variable Models

Myung Hwan Seo (Seoul Nat. U.)
with Bonsoo Koo, Masaya Takano (Monash)
Seojeong (Jay) Lee (SNU),

July 19, 2024
EcoSta 2024, Beijing

Background I

- A Structural impulse response (**IR**s), the reaction of a dynamic system in response to some external (policy) change, is a key causal parameter in empirical macroeconomics. IR

  1. Typically, it is defined through the Structural Vector Autoregression (SVAR), e.g.
  $$A_0 Y_t = A_1 Y_{t-1} + u_t,$$
  under diverse identifying assumptions on the matrix A_0 or $Eu_t u_t'$.
  2. The identification is also achieved by means of other external variables (or instruments).
  3. Sign restrictions result in **set identification**.
  4. It is often estimated by **Bayesian** approaches

Background II

- **Local Projection** and **Local Projection Instrumental Variable (LP-IV)** estimation are alternative popular approaches to SVAR.

  1. They are partial characterization of the system and intended to capture certain impulse responses only.
  2. They have not accommodated sign restrictions/set identification.
  3. In particular, it is not well understood what **Local Projection Instrumental Variable (LP-IV)** estimator identify under general conditions.

Background III

- Inference for set-identfied structural IR can be nonstandard and challenging.
    - **Uniformity** over a class of data distributions (Imbens and Manski (2004), Stoye (2009))
    - Boundary points of the identified set are possibly **nondifferentiable functions**: ($\Rightarrow$ no unbiased estimator, e.g. Hirano and Porter (2012))
    - The set may not be bounded.
- From the frequentist perspective (the set identified case):
    - Existing frequentist results: either conservative or only point-wise validity (Gafarov et al. (2018),Granziera et al. (2018))
    - Most Bayesian inference procedures can be invalid (Moon and Schorfheide (2012), Kitagawa et al. (2020)) and computationally demanding.
    - They are based on fullly specified **SVAR**. No work exists based on Local Projection (LP).

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.
- We establish certain set identification results for the case where multiple instruments are present and suitable sign restrictions are imposed.

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.
- We establish certain set identification results for the case where multiple instruments are present and suitable sign restrictions are imposed.
- We develop local asymptotic minimax inference:

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.
- We establish certain set identification results for the case where multiple instruments are present and suitable sign restrictions are imposed.
- We develop local asymptotic minimax inference:

  1. **shortest average length confidence interval** for IR maintaining uniform validity.

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.
- We establish certain set identification results for the case where multiple instruments are present and suitable sign restrictions are imposed.
- We develop local asymptotic minimax inference:

    1. **shortest average length confidence interval** for IR maintaining uniform validity.
    2. based on Local projection (Jordà, 2005) with instrumental variables (**LP-IV**): Stock and Watson (2018), Plagborg-Møller and Wolf (2021)

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.
- We establish certain set identification results for the case where multiple instruments are present and suitable sign restrictions are imposed.
- We develop local asymptotic minimax inference:

  1. **shortest average length confidence interval** for IR maintaining uniform validity.
  2. based on Local projection (Jordà, 2005) with instrumental variables (**LP-IV**): Stock and Watson (2018), Plagborg-Møller and Wolf (2021)
  3. overcomes the issues related to the nondifferentiability of boundary.

Overview of our Contribution

- We establish what an LP-IV estimand identifies in a reasonably general context.
- We establish certain set identification results for the case where multiple instruments are present and suitable sign restrictions are imposed.
- We develop local asymptotic minimax inference:

  1. **shortest average length confidence interval** for IR maintaining uniform validity.
  2. based on Local projection (Jordà, 2005) with instrumental variables (**LP-IV**): Stock and Watson (2018), Plagborg-Møller and Wolf (2021)
  3. overcomes the issues related to the nondifferentiability of boundary.
  4. computationally cheap.

Preview: LP-IV

Inference with external instruments

Stock and Watson (2018) consider identification of structural **IR**s to $\epsilon_{(1),t}$ in **LP-IV** (local projection with instrument variables) models with an instrument z_t for $\epsilon_{(1),t}$ where an LP-IV estimand is given by

$$\beta_{IV,(i,h)} = \mathbb{E}[y_{(i),t+h}z_t]/\mathbb{E}[y_{(1),t}z_t] \tag{1}$$

for $i = 1, \ldots, n$ and $h \geq 0$. Then,

$$\theta_{h,(i,1)} = \beta_{IV,(i,h)} \quad i = 1, \ldots, n \tag{2}$$

under the unit effect normalization and the following conditions:

Condition LP-IV

1. $\mathbb{E}\left[\epsilon_{(1),t}z_t\right] \neq 0$. (**relevance**)
2. $\mathbb{E}\left[\epsilon_{(2:n),t}z_t\right] = 0$. (**contemporaneous exogeneity**)
3. $\mathbb{E}\left[\epsilon_{t+j}z_t\right] = 0, \forall j \neq 0$. (**lead-lag exogeneity**)

Inference with external instruments: composite shock

- An instrument z_t may be correlated with multiple structural shocks (Jarociński and Karadi (2020), Braun and Brüggemann (2023), Koo et al. (2023)), in which case Condition LP-IV.2 (**contemporaneous exogeneity**) is violated.
- In this setup, Koo, Lee, Seo (2023), (**KLS23**), establish the LP-IV based set-identification of structural **IR**s under sign restrictions on a set of n_z instruments $\{z_t^{(k)}\}_{k=1}^{n_z}$, $n_z \in \mathbb{N}_+$.

Structural Vector Moving Average Model

$$\begin{pmatrix} x_t \\ \vdots \\ y_t \end{pmatrix} = \Theta(L) \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{m,t} \end{pmatrix}$$

$n \times 1$	$m \times 1$
observed	unobserved
endogenous	structural shocks

- $\varepsilon_{s,t}$'s are mutually uncorrelated, $E[\boldsymbol{\varepsilon}_t] = 0$, $E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] > 0$
- $\Theta(L) = \Theta_0 + \Theta_1 L + \Theta_2 L^2 + \cdots$, L is the lag operator
- Θ_h: $n \times m$ matrix of impulse responses whose elements are

$$\theta_{h,ys} \equiv E[y_{t+h} | \varepsilon_{s,t} = 1] - E[y_{t+h} | \varepsilon_{s,t} = 0]$$

An LP-IV Estimand May Not Be Meaningful

- Suppose x_t is the policy variable and the shock to x_t is $\xi_t = \varepsilon_{1,t} + \varepsilon_{2,t}$
- Under Assumption 1 and the unit effect normalization

$$\beta_{IV} = w_1\theta_{h,y1} + w_2\theta_{h,y2}$$

  where $w_s = Cov(z_t\varepsilon_{s,t})/(Cov(z_t\varepsilon_{1,t}) + Cov(z_t\varepsilon_{2,t}))$

- However, a weight can be negative
- Why is this a problem?

$$\begin{array}{ccccccc} \beta_h & = & w_1 & \theta_{h,y1} & + & w_2 & \theta_{h,y2} \\ \uparrow & & \uparrow & \uparrow & & \uparrow & \uparrow \\ 0 & & 2 & 1 & & -1 & 2 \end{array}$$

- "Negative weights" has been an important issue in econometrics in recent years, e.g.,
  - TWFE: de Chaisemartin and D'Haultfoeuill (2020, AER)
  - 2SLS: Mogstad, Torgovitsky, and Walters (2021, AER)

Two LP-IV estimands can yield an Identified Set

- Two instruments: z_t^A and z_t^B
- Two separate LP-IV estimands: β_h^A and β_h^B
- Intersection gives one of the following identified sets:

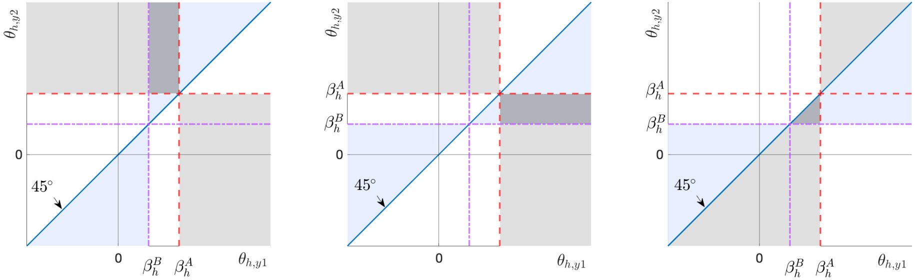
$$UL = \{\theta_{h,ys} : \theta_{h,ys} < \min(\beta_h^A, \beta_h^B) \text{ or } \max(\beta_h^A, \beta_h^B) < \theta_{h,ys}\} \tag{3}$$

$$U = \{\theta_{h,ys} : \theta_{h,ys} < \min(\beta_h^A, \beta_h^B)\} \tag{4}$$

$$L = \{\theta_{h,ys} : \max(\beta_h^A, \beta_h^B) < \theta_{h,ys}\} \tag{5}$$

$$T = \{\theta_{h,ys} : \min(\beta_h^A, \beta_h^B) < \theta_{h,ys} < \max(\beta_h^A, \beta_h^B)\} \tag{6}$$

- T is the most informative, UL is the least



(d) $\Theta_{++}^A \cap \Theta_{+-}^B,\ \beta_h^A > \beta_h^B$ (e) $\Theta_{++}^A \cap \Theta_{-+}^B,\ \beta_h^A > \beta_h^B$ (f) $\Theta_{+-}^A \cap \Theta_{-+}^B,\ \beta_h^A > \beta_h^B$

- Two IVs with different signs ($+-$ vs $-+$) gives the informative bounds for both $\theta_{h,y1}$ and $\theta_{h,y2}$

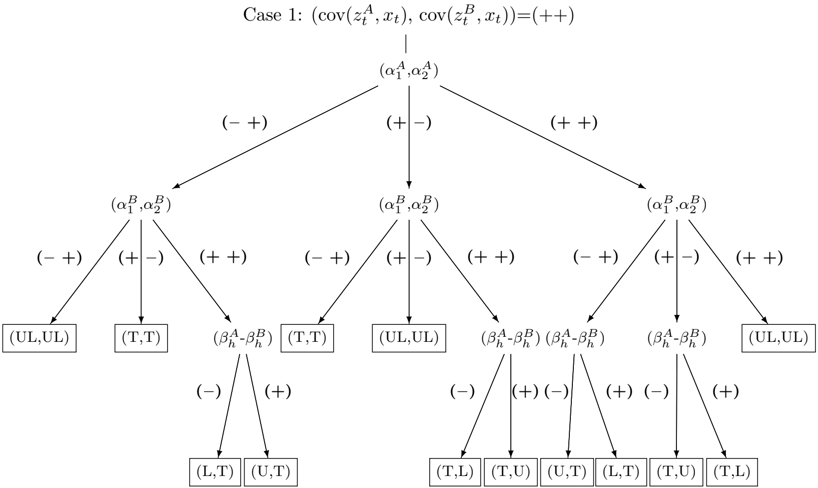
LP-IV as a Bound of the Identified Set

Proposition 1

Suppose that random variables z_t^A and z_t^B satisfy Assumption 1. We assume that $cov(z_t^A, x_t) > 0$, $cov(z_t^B, x_t) > 0$, and $\beta_h^A > \beta_h^B$ without loss of generality.

1. *If $\Theta^A_{++} \cap \Theta^B_{++}$ or $\Theta^A_{+-} \cap \Theta^B_{+-}$ or $\Theta^A_{-+} \cap \Theta^B_{-+}$, then $\theta_{h,y1} \in UL$ and $\theta_{h,y2} \in UL$.*
2. *If $\Theta^A_{++} \cap \Theta^B_{+-}$, then $\theta_{h,y1} \in T$ and $\theta_{h,y2} \in L$.*
3. *If $\Theta^A_{+-} \cap \Theta^B_{++}$, then $\theta_{h,y1} \in T$ and $\theta_{h,y2} \in U$.*
4. *If $\Theta^A_{++} \cap \Theta^B_{-+}$, then $\theta_{h,y1} \in L$ and $\theta_{h,y2} \in T$.*
5. *If $\Theta^A_{-+} \cap \Theta^B_{++}$, then $\theta_{h,y1} \in U$ and $\theta_{h,y2} \in T$.*
6. *If $\Theta^A_{+-} \cap \Theta^B_{-+}$ or $\Theta^A_{-+} \cap \Theta^B_{+-}$, then $\theta_{h,y1} \in T$ and $\theta_{h,y2} \in T$.*

Reference for Practitioners



Example: Jarociński and Karadi (2020)

- The US Federal Open Market Committee (FOMC) announcements are usually considered as monetary policy shocks
- Observes that the interest rate and stock price surprises often move in the same direction upon the FOMC announcements
- FOMC announcement:

$$\xi_t = \varepsilon_{mp,t} + \varepsilon_{cb,t}$$

- $\varepsilon_{mp,t}$: pure monetary policy shock
- $\varepsilon_{cb,t}$: the central bank information shock

IV and Sign Restrictions

- High-frequency financial market surprises (the change between 10 minutes before and 20 minutes after the announcements)
- Two IVs: z_t^{ff} and z_t^{sp}
- z_t^{ff}: high-frequency surprise in the fed funds futures
- z_t^{sp}: high-frequency surprise in the stock price
- Sign restrictions are

$$
\begin{array}{ll}
z_t^{ff}: & Cov(z_t^{ff}, \varepsilon_{mp,t}) > 0, \quad Cov(z_t^{ff}, \varepsilon_{cb,t}) > 0 \quad \Rightarrow \Theta_{++}^{ff} \\
z_t^{sp}: & Cov(z_t^{sp}, \varepsilon_{mp,t}) < 0, \quad Cov(z_t^{sp}, \varepsilon_{cb,t}) > 0 \quad \Rightarrow \Theta_{-+}^{sp}
\end{array}
$$

Identified Set for Componentwise IR

- Recall

$$\beta_h^j = w_{mp}^j \theta_{h,mp} + w_{cb}^j \theta_{h,cb}$$

where

$$w_s^j = \frac{Cov(z_t^j, \varepsilon_{s,t})}{Cov(z_t^j, x_t)}$$

for $j = ff, sp$ and $s = mp, cb$

- $\theta_{h,mp}$: response of y_{t+h} to the unit change in the pure monetary shock (corresponding to 25BP change in x_t)
- $\theta_{h,cb}$: response of y_{t+h} to the unit change in the central bank information shock (corresponding to 25BP change in x_t)
- Obtain the identified sets L and T for each component IR by $\Theta_{++}^{ff} \cap \Theta_{-+}^{sp}$.

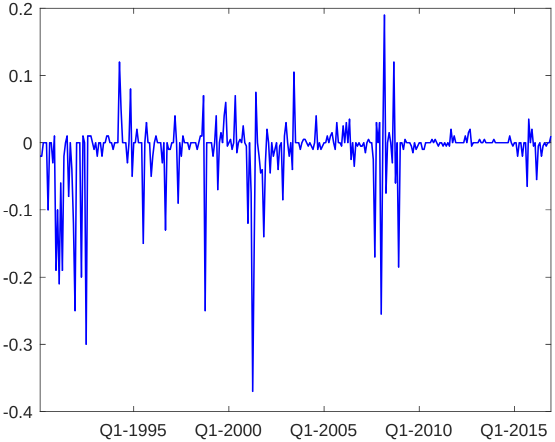
Data and Model

- Econometric model:

$$y_{t+h} = \mu_h + \beta_h x_t + \boldsymbol{\phi}_h(L)'\boldsymbol{R}_{t-1} + u_{t+h}$$

- y_t: macro variable of interest
- x_t: monthly average of the 1-year Treasury yield
- $\boldsymbol{R}_t$: all the macro var including y_t, x_t, and z_t
- $\boldsymbol{\phi}_h(L)$: a coeff vector of polynomial in the lag operator of order 12
- β_h: Responses to 25BP increase in 1-year government bond yield
- $\widehat{\beta}_h^{ff}$, $\widehat{\beta}_h^{sp}$: LP-IV est using z_t^{ff} and z_t^{sp} one at a time, respectively

high-frequency surprises in the fed funds futures



Source: Gertler and Karadi (2015)

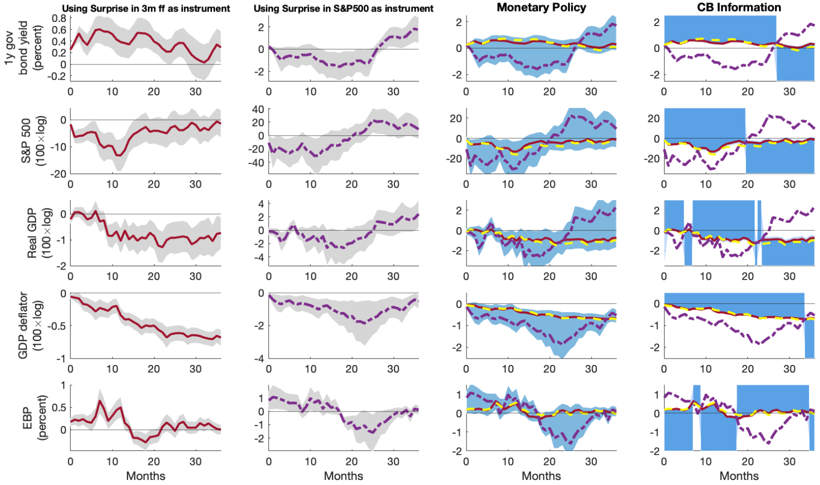


Figure: Responses to 25BP Increase in One-year Government Bond Yield

go back

Summary

- IR analysis using LP-IV when the macro shock is composite
- LP-IV estimand is an affine combination of componentwise IR (non-negative weights under the same-sign condition)
- Main intuition is that the IV may be heterogeneously correlated with the macro shock elements, which generates heterogeneous treatment
- LP-IV estimand should be carefully interpreted as a structural IR
- Simple to implement identification strategies for componentwise IR are proposed

Besides, it is Not Okay to use 2SLS

- LP-2SLS or LP-GMM with multiple instruments, e.g., Stock and Watson (2018), Ramey and Zubairy (2018), does not work.
- Why? Andrews (2019) shows 2SLS is a linear combination of ind. IV estimand
- 2SLS can be non-causal even when the causal LP-IV estimands are combined
- An example is given in the paper
- Additionally, the overidentifying restrictions test is not recommended

Optimal Confidence Interval

Inference for an interval identified parameter

For a univariate parameter of interest $\theta \in \Omega_\theta \subset \mathbb{R}$, suppose the identified set Θ_0 for θ given the data distribution P is given by

$$\Theta_0(P) = [\theta_l, \theta_u] \tag{7}$$

where a pair of boundary points (θ_l, θ_u) is a functional of some reduced-form parameter $\beta(P) \in \Omega_\beta \subset \mathbb{R}^{n_\beta}$:

$$\Gamma(\beta(P)) = \begin{bmatrix} \theta_l \\ \theta_u \end{bmatrix} \tag{8}$$

for some mapping Γ from the parameter space Ω_β of β to the space of (θ_l, θ_u):

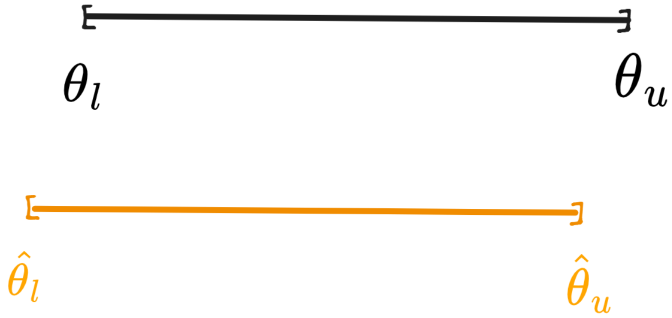
$$\{(\theta_l, \theta_u)^\top \in \Omega_\theta \times \Omega_\theta \, : \, \theta_l \leq \theta_u\} \subset \mathbb{R}^2. \tag{9}$$

Optimal Confidence regions for the parameter

Our goail is the construction of minimax optimal confidence regions (**CR**s).
Two notions of CRs under set-identification:

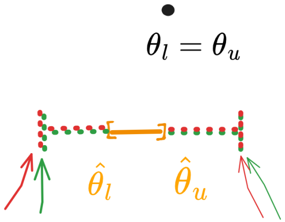
1. **CR** for the identified set Θ_0
2. **CR** for the "true" value θ_0, cf. Imbens and Manski (2004).

Estimated identified set $\hat{\Theta} = [\hat{\theta}_l, \hat{\theta}_u]$



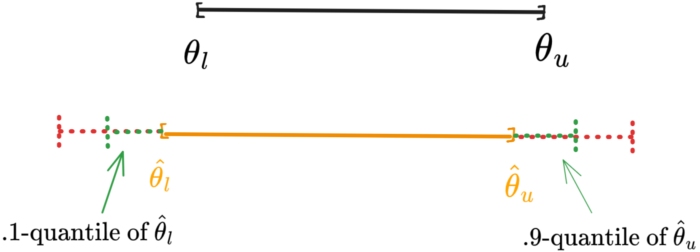
$\mathcal{CR}_{1-\alpha}$: **CR** for the point-identified case

with size .9



$\mathcal{CR}_{1-\alpha}$: **CR** for the set-identified case

with size .9



Nonstandard inference: 1. Uniformity

In practice, it is typically unknown a priori whether θ is set- or point-identified.

- Ruling out point-identification leads to under-coverage when θ is (or is close to be) indeed point-identified.

As pointed out by Imbens and Manski (2004), a **CR** for the parameter θ needs to be uniformly valid over a given class $\mathcal{P}$ of data distributions P:

$$\inf_{P \in \mathcal{P}} \inf_{\theta \in \Theta_0(P)} \Pr_P(\theta \in \mathcal{CR}_{1-\alpha}) \geq 1 - \alpha. \tag{10}$$

to ensure finite-sample validity under (near) point-identification.

Nonstandard inference: 2. Nondifferentiability of Γ

Recall

$$\begin{bmatrix}\theta_l \\ \theta_u\end{bmatrix} = \Gamma(\beta(P)) \tag{11}$$

for some functional Γ of a finite-dimensional reduced form parameter β.
Typically, a $\sqrt{T}$-consistent and asymptotically normal estimator $\hat{\beta}$ of $\beta(P)$ is available:

$$\sqrt{T}(\hat{\beta} - \beta(P)) \xrightarrow{d} \mathcal{N}[\mathbf{0}, \Sigma] \tag{12}$$

where Σ is positive definite.
Let $(\hat{\theta}_l, \hat{\theta_u})$ be a (naive) plug-in estimator for (θ_l, θ_u) defined as

$$\begin{bmatrix}\hat{\theta}_l \\ \hat{\theta}_u\end{bmatrix} = \Gamma(\hat{\beta}) \tag{13}$$

Nonstandard inference: 2. Nondifferentiability of Γ

If Γ is differentiable, $(\hat{\theta}_l, \hat{\theta})_u$ is $\sqrt{T}$-consistent and asymptotically normal by the continuous mapping theorem.

However, for set-identified IRs, Γ may be only continuous but not differentiable at $\beta(P)$, in which case Hirano and Porter (2012) show

- There exists no unbiased estimator for (θ_l, θ_u).
- There exists no regular estimator for (θ_l, θ_u).

    - Asymptotic properties of estimators are not invariant to $\sqrt{T}-$ local pertubation of P.
    - e.g. $\max(\beta), \min(\beta)$ are non-differentiable on the line that $\beta_1 = \beta_2$.

The asymptotic distribution of an estimator for (θ_l, θ_u) is not normal in general.

Nonstandard inference: 2. Remarks on nondifferentiability issues

If Γ is not differentiable, a "good" estimator $\check{\beta}$ (e.g., semiparametric efficient) for β does not necessarily lead to a "good" plug-in estimator $\Gamma(\check{\beta})$.

Then, what are desirable properties of $\mathcal{CR}_{1-\alpha}$ for θ we wish to achieve?

- $\mathcal{CR}_{1-\alpha}$ has uniform level of $1-\alpha$ over some family of $\sqrt{T}$-local pertubations of P.
- The average length of $\mathcal{CR}_{1-\alpha}$ is as small as possible while maintaining the uniform level.

We formulate this problem as a minimization problem of a local asymptotic minimax risk, Hájek (1972), for $\mathcal{CR}_{1-\alpha}$.

Local asymptotic minimax inference

- Song (2014) and Fang (2014) study local asymptotic minimax optimality in point estimation problems of a univariate parameter when Γ is possibly nondifferentiable.
- We consider local asymptotic minimax problems for confidence regions, which are random sets which depend on a multivariate parameter vector and nonlinear constraints.

Element-wise additive corrections

- Given a $\sqrt{T}$-consistent and asymptotically normal estimator $\hat{\beta}$, we consider a plug-in type estimator $\Gamma(\tilde{\beta})$ with element-wise corrections where

$$\tilde{\beta} = \hat{\beta} + \tilde{c}/\sqrt{T} \tag{14}$$

where $\tilde{c}$ is a sequence of $\mathbb{R}^{n_\beta} \times 1$ random vectors which converges in probability to some nonrandom limit $c \in \mathbb{R}^{n_\beta}$.

  - Optimality of element-wise corrections of the form (14) is known for some class of loss functions (Van der Vaart (1992), Song (2014), Fang (2014)).

on Bayesian inference for set-identified models

- For point-identified models, a Bayesian credible set is an asymptotically valid confidence set from the frequentist view by the Bernstein–von Mises theorem.
- For **set-identified** models, it is not (Moon and Schorfheide (2012)).
    - The prior belief is not asymptotically negligible.
    - "asymptotically Bayesian highest-posterior- density sets exclude parts of the estimatedm identified set, whereas it is well known that frequentist confidence sets extend beyond the boundaries of the estimated identified set."
    - Many of existing empirical works belong to this case.

- Giacomini and Kitagawa (2021) propose a multiple-prior Bayesian inference approach which provides a robust Bayesian credible set for the identified set with the correct frequentist coverage probability.

on Bayesian inference for set-identified models

- For point-identified models, a Bayesian credible set is an asymptotically valid confidence set from the frequentist view by the Bernstein–von Mises theorem.
- For **set-identified** models, it is not (Moon and Schorfheide (2012)).
  - The prior belief is not asymptotically negligible.
  - "asymptotically Bayesian highest-posterior- density sets exclude parts of the estimatedm identified set, whereas it is well known that frequentist confidence sets extend beyond the boundaries of the estimated identified set."
  - Many of existing empirical works belong to this case.
- Giacomini and Kitagawa (2021) propose a multiple-prior Bayesian inference approach which provides a robust Bayesian credible set for the identified set with the correct frequentist coverage probability.
  - However, GK relies crucially on the assumption that Γ is differentiable and the set is bounded.

Local Asymptotically Normal (LAN) limit experiment

We consider the local drifting sequence $\{\beta_T(\mathfrak{h})\}$ of the reduced form regular parameter defined as

$$\beta_T(\mathfrak{h}) = \beta + \mathfrak{h}/\sqrt{T}, \quad \mathfrak{h} \in \mathcal{H} \tag{15}$$

where $\mathcal{H} \subset \mathbb{R}^2$.

Assumption 1 (Regular estimator $\hat{\beta}$ of β)

1. *$\hat{\beta}$ is a regular asymptotically linear estimator of $\beta(P_0)$:*

$$\sqrt{T}(\hat{\beta} - \beta_T(\mathfrak{h})) \xrightarrow{P_{\mathfrak{h},T}} \mathcal{N}[\mathbf{0}, \Sigma] \quad \forall \mathfrak{h} \in \mathcal{H} \tag{16}$$

   where Σ is positive definite.

2. *The estimator $\hat{\Sigma}$ converges in probability to Σ uniformly over $\mathfrak{h} \in \mathcal{H}$.*

Local Asymptotic Minimax Optimal **CR**

We are interested in the local asymptotic minimax optimal confidence region for θ at a level $1-\alpha$ in terms of the minimum average length. More specifically, we construct a random set $\mathcal{CR}_{1-\alpha}$ such that

$$\mathcal{R}(\mathcal{CR}_{1-\alpha}) = \sup_{b\in[0,\infty)} \liminf_{T\to\infty} \sup_{\mathfrak{h}\in\mathcal{H}_b} \mathbb{E}_{T,\mathfrak{h}}\left[\sqrt{T}\mathrm{vol}\left(\mathcal{CR}_{1-\alpha}\backslash\Theta_0(P_{T,\mathfrak{h}})\right)\right] \tag{17}$$

subject to

$$\sup_{b\in[0,\infty)} \limsup_{T\to\infty} \sup_{\mathfrak{h}\in\mathcal{H}_b} \sup_{\theta\in\Theta_0(P_{T,\mathfrak{h}})} \Pr_{T,\mathfrak{h}}\left(\theta\notin\mathcal{CR}_{1-\alpha}\right) \leq a. \tag{18}$$

where $\mathcal{H}_b = \{\mathfrak{h}\in\mathcal{H} : \|\mathfrak{h}\| \leq b\}$.

Leading case

Here, we focus

1. on the class of

$$\begin{aligned} \Theta_0(P) &= [\min(\beta_{(1)}(P), \beta_{(2)}(P)), \max(\beta_{(1)}(P), \beta_{(2)}(P))] \\ &= [\beta_{(1)}(P), \beta_{(2)}(P)] \cup [\beta_{(2)}(P), \beta_{(1)}(P)] \end{aligned} \tag{19}$$

2. on the near-identified case where both $\beta_{(1)}$ and $\beta_{(2)}$ is in a $\sqrt{T}$-local neighborhood of θ, i.e. $\beta = (\theta, \theta)^\top$ in (15).

Then,

$$\Theta_0(P_{T,\mathfrak{h}}) = \begin{cases} [\beta_{T,(1)}(\mathfrak{h}_{(1)}), \beta_{T,(2)}(\mathfrak{h}_{(2)})] & \text{if } \mathfrak{h}_{(1)} \leq \mathfrak{h}_{(2)} \\ [\beta_{T,(2)}(\mathfrak{h}_{(2)}), \beta_{T,(1)}(\mathfrak{h}_{(1)})] & \text{if } \mathfrak{h}_{(1)} > \mathfrak{h}_{(2)} \end{cases}$$

Class of CR's

- we restrict our attention to a plug-in type estimator with element-wise additive correction terms:

$$\begin{aligned} \mathcal{CR}_{1-\alpha} \quad = \quad & [\min(\hat{\beta}_{(1)} + \tilde{c}_{l,(1)}/\sqrt{T}, \hat{\beta}_{(2)} + \tilde{c}_{l,(2)}/\sqrt{T}), \\ & \max(\hat{\beta}_{(1)} + \tilde{c}_{u,(1)}/\sqrt{T}, \hat{\beta}_{(2)} + \tilde{c}_{u,(2)}/\sqrt{T})] \end{aligned} \tag{20}$$

- in line with Chernozhukov, Lee, Rosen (2013) in the sense that we do not construct the critical values based on the distributions of the boundary estimators, like $[\min(\hat{\beta}) + c_1/\sqrt{T}, \max(\hat{\beta}) + c_2/\sqrt{T}]$.

- optimality of element-wise corrections of this form is known for some class of loss functions (Van der Vaart (1992), Song (2014), Fang (2014)) in case of the point estimation.

Class of CR's

- we restrict our attention to a plug-in type estimator with element-wise additive correction terms:

$$\mathcal{CR}_{1-\alpha} = [\min(\hat{\beta}_{(1)} + \tilde{c}_{l,(1)}/\sqrt{T}, \hat{\beta}_{(2)} + \tilde{c}_{l,(2)}/\sqrt{T}), \max(\hat{\beta}_{(1)} + \tilde{c}_{u,(1)}/\sqrt{T}, \hat{\beta}_{(2)} + \tilde{c}_{u,(2)}/\sqrt{T})] \quad (20)$$

- in line with Chernozhukov, Lee, Rosen (2013) in the sense that we do not construct the critical values based on the distributions of the boundary estimators, like $[\min(\hat{\beta}) + c_1/\sqrt{T}, \max(\hat{\beta}) + c_2/\sqrt{T}]$.
- optimality of element-wise corrections of this form is known for some class of loss functions (Van der Vaart (1992), Song (2014), Fang (2014)) in case of the point estimation.
- However, the minimax optimization problem (20) can be **computationally prohibitive.**

Under this limit experiment

The risk $\ddot{\mathcal{R}}_\alpha(c\,;\,\hat{\beta})$ simplifies to

$$\begin{aligned}\ddot{\mathcal{R}}_{1-\alpha}(c\,;\,\hat{\beta}) \quad &= \quad \mathbb{E}\Big[\max(Z_{(1)}(\Sigma)+c_{u,(1)}, Z_{(2)}(\Sigma)+c_{u,(2)})\Big] \\ &\quad -\mathbb{E}\Big[\min(Z_{(1)}(\Sigma)+c_{l,(1)}, Z_{(2)}(\Sigma)+c_{l,(2)})\Big] \end{aligned} \tag{21}$$

subject to

$$\begin{aligned}\inf_{v\in\{\tilde{v}\in\mathbb{R}^2\,;\,\tilde{v}_{(1)}\tilde{v}_{(2)}\leq 0\}} \Pr\Big(&\max(Z_{(1)}(\Sigma)+c_{u,(1)}+v_{(1)}, Z_{(2)}(\Sigma)+c_{u,(2)}+v_{(2)}) \geq 0 \\ &\geq \min(Z_{(1)}(\Sigma)+c_{l,(1)}+v_{(1)}, Z_{(2)}(\Sigma)+c_{l,(2)}+v_{(2)})\Big) \geq 1-\alpha. \end{aligned} \tag{22}$$

where

$$\begin{bmatrix} Z_{(1)}(\Sigma) \\ Z_{(2)}(\Sigma) \end{bmatrix} \sim \mathcal{N}\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma\right] \tag{23}$$

Theorem: Optimal CI for θ_0

- Let $c_{l,(1)} = -c_{u,(1)}, c_{u,(2)} = -c_{l,(2)}$ and

$$(c^*_{(1)}, c^*_{(2)}) = \arg\min c_{u,(1)} - c_{l,(2)} \quad \text{s.t. (22) holds.}$$

- Then, the unique minimizer $c' = ((c'_u)^\top, (c'_l)^\top)^\top$ of the local asymptotic minimax risk $\mathcal{R}(\mathcal{CR}_{1-\alpha})$ over $c \in \mathbb{R}^4$ given $\hat{\beta}$ is characterized as

1. If $\sigma_1 = \sigma_2$,

$$c' = (c^*_{(1)}, c^*_{(2)}, -c^*_{(1)}, -c^*_{(2)}). \tag{24}$$

2. If $\sigma_1 < \sigma_2$, there exist some $\eta^* > 0$ and $\lambda_{\eta^*} \in (0,1)$ such that

$$c' = (c^*_{(1)} + \eta^*, c^*_{(2)} - \lambda_{\eta^*}\eta^*, -c^*_{(1)} - \eta^*, -c^*_{(2)} + \lambda_{\eta^*}\eta^*)^\top. \tag{25}$$

3. If $\sigma_1 > \sigma_2$, there exist some $\eta^{**} \geq > 0$ and $\lambda_{\eta^{**}} \in (0,1)$ such that

$$c' = (c^*_{(1)} - \lambda_{\eta^{**}}\eta^{**}, c^*_{(2)} + \eta^{**}, -c^*_{(1)} + \lambda_{\eta^{**}}\eta^{**}, -c^*_{(2)} - \eta^*)^\top. \tag{26}$$

Monte Carlo Simulation

Monte Carlo experiments: Design (Instruments)

- 2 instruments $z_t^{(1)}$ and $z_t^{(2)}$ correlated with $\epsilon_{j,t}$, $j = 1, 2$:

$$\begin{bmatrix} z_t^{(1)} \\ z_t^{(2)} \end{bmatrix} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \tag{27}$$

where $u_t \overset{\text{i.i.d.,}}{\sim} \mathcal{N}\left[\mathbf{0}, I_2\right]$ and $u_t \perp \epsilon_t$.

  - $\alpha_{1,1} > 0$, $\alpha_{1,2} > 0$ and $\alpha_{2,1} > 0$, $\alpha_{2,2} < 0$, $\alpha_{2,1} + \alpha_{2,2} > 0$ as in the example
  - The LP-IV estimators

$$\hat{\beta}_{(1),h} = \frac{\mathbb{E}_T[z_t^{(1)} y_{t+h}]}{\mathbb{E}_T[z_t^{(1)} x_t]}, \quad \hat{\beta}_{(2),h} = \frac{\mathbb{E}_T[z_t^{(2)} y_{t+h}]}{\mathbb{E}_T[z_t^{(2)} x_t]}. \tag{28}$$

Monte Carlo experiments: Design

- The lag-order of SVMA: $\bar{L} = 16$
- The sample size $T = 500$
- Coefficient for $(z_t^{(1)}, z_t^{(2)})$:

$$\begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} = \begin{bmatrix} 2 & .4 \\ 1.2 & -.5 \end{bmatrix}$$

- For each replication, we construct CRs for each $\theta_{1,h}$ given $\{y_t, x_t, z_t^{(1)}, z_t^{(2)}\}$ with level .9 and compute empirical coverage probabilities and the average values of the upper and lower bounds of the CR (# of replications = 300).

Monte Carlo experiments: Comparison with a naive estimator

We compare a local asymptotic minimax CR with the following naive CR:

$$\left[\hat{\beta}_l - T^{-1/2}\hat{\sigma}_l q_{1-\alpha_1}, \hat{\beta}_u + T^{-1/2}\hat{\sigma}_u q_{1-\alpha_2}\right]$$

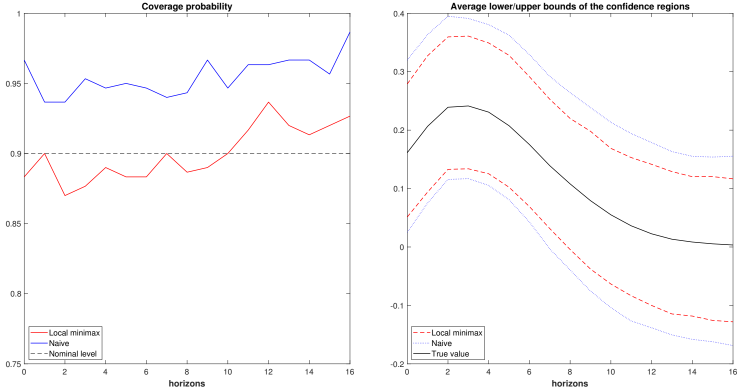
where $l \in \{1, 2\}$ is the index of the smaller of $\{\hat{\beta}_{(1)}, \hat{\beta}_{(2)}\}$:

$$\hat{\beta}_l = \min\left(\hat{\beta}_{(1)}, \hat{\beta}_{(2)}\right), \; \hat{\beta}_u = \max\left(\hat{\beta}_{(1)}, \hat{\beta}_{(2)}\right),$$

$(\hat{\sigma}_l, \hat{\sigma}_u)$ are corresponding estimated asymptotic standard errors.
A pair $(\alpha_1, \alpha_2) \in (0, 1)^2$ satisfies $(1 - \alpha_1)(1 - \alpha_2) = 1 - \alpha$ and is chosen to minimize the length of the CR: $\hat{\sigma}_u q_{1-\alpha_1} + \hat{\sigma}_l q_{1-\alpha_1}$.

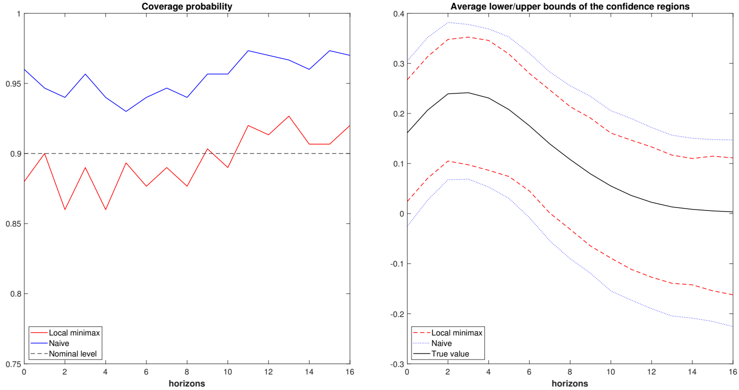
Coverage and Length of CI's

$\delta = 0$



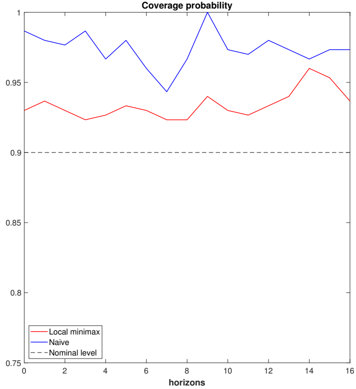
Coverage and Length of CI's

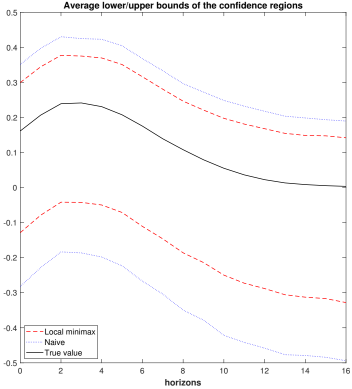
$\delta = 1$



Coverage and Length of CI's

$\delta = 10$





Conclusion

- We showed that an LP-IV estimand may not be causal.

Conclusion

- We showed that an LP-IV estimand may not be causal.
- We showed that combination of multiple LP-IV estimands forms an identified set when suitable sign restrictions can be imposed.

Conclusion

- We showed that an LP-IV estimand may not be causal.
- We showed that combination of multiple LP-IV estimands forms an identified set when suitable sign restrictions can be imposed.
- We developed local asymptotic minimax inference:

Conclusion

- We showed that an LP-IV estimand may not be causal.
- We showed that combination of multiple LP-IV estimands forms an identified set when suitable sign restrictions can be imposed.
- We developed local asymptotic minimax inference:

    1. based on simple LP-IV estimands

Conclusion

- We showed that an LP-IV estimand may not be causal.
- We showed that combination of multiple LP-IV estimands forms an identified set when suitable sign restrictions can be imposed.
- We developed local asymptotic minimax inference:
  1. based on simple LP-IV estimands
  2. **shortest average length confidence interval** for IR maintaining uniform validity.

Conclusion

- We showed that an LP-IV estimand may not be causal.
- We showed that combination of multiple LP-IV estimands forms an identified set when suitable sign restrictions can be imposed.
- We developed local asymptotic minimax inference:

  1. based on simple LP-IV estimands
  2. **shortest average length confidence interval** for IR maintaining uniform validity.
  3. **computationally fast**.

Thank you for your attention!

References I

Braun, R. and Brüggemann, R. (2023). Identification of SVAR Models by Combining Sign Restrictions With External Instruments. *Journal of Business & Economic Statistics Economic Statistics*, 41(4):1077–1089.

Fang, Z. (2014). Optimal plug-in estimators of directionally differentiable functionals. *Unpublished Manuscript*.

Gafarov, B., Meier, M., and Olea, J. L. M. (2018). Delta-method inference for a class of set-identified SVARs. *Journal of Econometrics*, 203(2):316–327.

Giacomini, R. and Kitagawa, T. (2021). Robust Bayesian inference for set-identified models. *Econometrica*, 89(4):1519–1556.

Granziera, E., Moon, H. R., and Schorfheide, F. (2018). Inference for VARs identified with sign restrictions. *Quantitative Economics*, 9(3):1087–1121.

Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 175–194.

Hirano, K. and Porter, J. R. (2012). Impossibility Results for Nondifferentiable Functionals. *Econometrica*, 80(4):1769–1790.

References II

Imbens, G. W. and Manski, C. F. (2004). Confidence Intervals for Partially Identified Parameters. *Econometrica*, 72(6):1845–1857.

Jarociński, M. and Karadi, P. (2020). Deconstructing Monetary Policy Surprises—The Role of Information Shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Kitagawa, T., Montiel Olea, J. L., Payne, J., and Velez, A. (2020). Posterior distribution of nondifferentiable functions. *Journal of Econometrics*, 217(1):161–175.

Koo, B., Lee, S., and Seo, M. H. (2023). What Impulse Response Do Instrumental Variables Identify?

Moon, H. R. and Schorfheide, F. (2012). Bayesian and Frequentist Inference in Partially Identified Models. *Econometrica*, 80(2):755–782.

Song, K. (2014). Local asymptotic minimax estimation of nonregular parameters with translation-scale equivariant maps. *Journal of Multivariate Analysis*, 125:136–158.

References III

Stock, J. H. and Watson, M. W. (2018). Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *The Economic Journal*, 128(610):917–948.

Stoye, J. (2009). More on confidence intervals for partially identified parameters. *Econometrica*, 77(4):1299–1315.

Van der Vaart, AW. (1992). Asymptotic linearity of minimax estimators. *Statistica neerlandica*, 46(2-3):179–194.