A New Semiparametric Approach for Nonstandard Econometric Problems

Bo Zhou

Tilburg University

November 10, 2016

Motivation

- Consider the semiparametric model specified with given ϕ

$$y_t = \phi\big(y_{t-1}, x_t, z_t; \theta\big) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim f \tag{1}$$

  where θ is the low-dim. parametric part of interest, while f is the infinite-dim. nuisance nonparametric part.

- This paper is motivated by:
  - (I) Considerable information contained in non-Gaussian density f
  - (II) Commonly-used Gaussian-based methods cannot use these information
  - (III) The traditional semiparametric approach does not work in nonstandard econometric problems

- A NEW semiparametric approach is needed for these nonstandard problems!

Motivation

- Consider the semiparametric model specified with given ϕ

$$y_t = \phi\big(y_{t-1}, x_t, z_t; \theta\big) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim f \tag{1}$$

  where θ is the low-dim. parametric part of interest, while f is the infinite-dim. nuisance nonparametric part.

- This paper is motivated by:
  - (I) Considerable information contained in non-Gaussian density f
  - (II) Commonly-used Gaussian-based methods cannot use these information
  - (III) The traditional semiparametric approach does not work in nonstandard econometric problems

- A NEW semiparametric approach is needed for these nonstandard problems!

Motivation

- Consider the semiparametric model specified with given ϕ

$$y_t = \phi\big(y_{t-1}, x_t, z_t; \theta\big) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim f \tag{1}$$

  where θ is the low-dim. parametric part of interest, while f is the infinite-dim. nuisance nonparametric part.

- This paper is motivated by:
  (I) Considerable information contained in non-Gaussian density f
  (II) Commonly-used Gaussian-based methods cannot use these information
  (III) The traditional semiparametric approach does not work in nonstandard econometric problems

- A NEW semiparametric approach is needed for these nonstandard problems!

Motivation

- Consider the semiparametric model specified with given ϕ

$$y_t = \phi\big(y_{t-1}, x_t, z_t; \theta\big) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim f \tag{1}$$

  where θ is the low-dim. parametric part of interest, while f is the infinite-dim. nuisance nonparametric part.

- This paper is motivated by:

  (I) Considerable information contained in non-Gaussian density f

  (II) Commonly-used Gaussian-based methods cannot use these information

  (III) The traditional semiparametric approach does not work in nonstandard econometric problems

- A NEW semiparametric approach is needed for these nonstandard problems!

Motivation

- Consider the semiparametric model specified with given ϕ

$$y_t = \phi\big(y_{t-1}, x_t, z_t; \theta\big) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim f \tag{1}$$

  where θ is the low-dim. parametric part of interest, while f is the infinite-dim. nuisance nonparametric part.

- This paper is motivated by:
  (I) Considerable information contained in non-Gaussian density f

  (II) Commonly-used Gaussian-based methods cannot use these information

  (III) The traditional semiparametric approach does not work in nonstandard econometric problems

- A NEW semiparametric approach is needed for these nonstandard problems!

Motivation

- Consider the semiparametric model specified with given ϕ

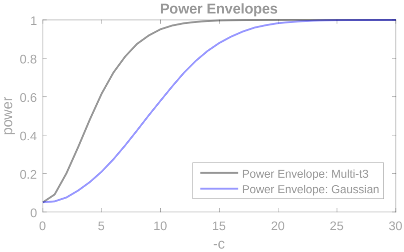
$$y_t = \phi\big(y_{t-1}, x_t, z_t; \theta\big) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim f \tag{1}$$

where θ is the low-dim. parametric part of interest, while f is the infinite-dim. nuisance nonparametric part.

- This paper is motivated by:
  - (I) Considerable information contained in non-Gaussian density f
  - (II) Commonly-used Gaussian-based methods cannot use these information
  - (III) The traditional semiparametric approach does not work in nonstandard econometric problems
- A NEW semiparametric approach is needed for these nonstandard problems!

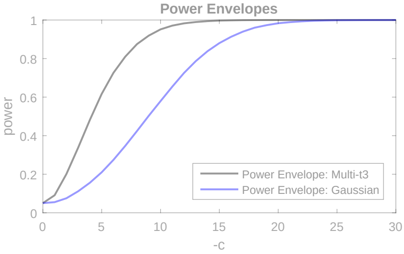
Point (I): Considerable information contained in non-Gaussian density f

- Information for what? — Efficiency of statistical inferences
  - Point estimation: lower standard deviations
  - Testing hypothesis: higher testing powers
  - Both $\Longrightarrow$ narrower confidence intervals
- How much efficiency can we gain from non-Gaussian f? — Take the cointegration case as an example



Point (I): Considerable information contained in non-Gaussian density f

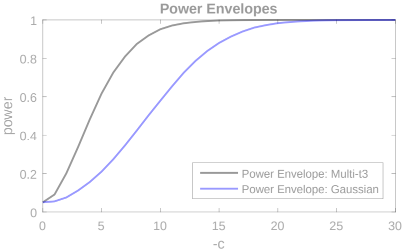
- Information for what? — Efficiency of statistical inferences
    - Point estimation: lower standard deviations
    - Testing hypothesis: higher testing powers
    - Both $\Longrightarrow$ narrower confidence intervals
- How much efficiency can we gain from non-Gaussian f? — Take the cointegration case as an example



Introduction

Point (I): Considerable information contained in non-Gaussian density f

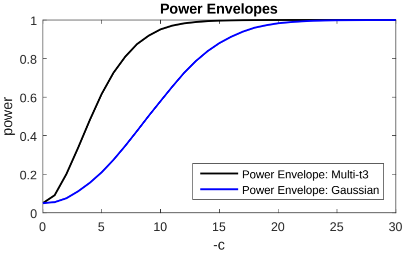
- Information for what? — Efficiency of statistical inferences
  - Point estimation: lower standard deviations
  - Testing hypothesis: higher testing powers
  - Both $\Longrightarrow$ narrower confidence intervals
- How much efficiency can we gain from non-Gaussian f? — Take the cointegration case as an example



Introduction

Point (I): Considerable information contained in non-Gaussian density f

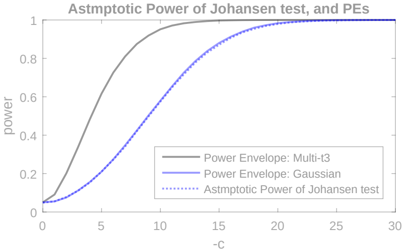
- Information for what? — Efficiency of statistical inferences
  - Point estimation: lower standard deviations
  - Testing hypothesis: higher testing powers
  - Both $\Longrightarrow$ narrower confidence intervals
- How much efficiency can we gain from non-Gaussian f? — Take the cointegration case as an example



Introduction

Point (II): Commonly-used Gaussian-based methods cannot use these information

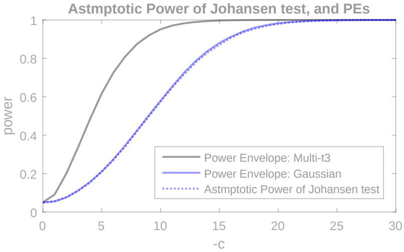
- What are Gaussian-based methods? — Muller (2011)
  - Examples: OLS and t-test for linear regression; the Dickey-Fuller test, the ERS test for unit root; the Johansen test for cointegration; the AR test, the LM test, and the CLR test for weak instruments ...
  - Step I: assume i.i.d ε_t, and f Gaussian $\longrightarrow$ test statistics
  - Step II: show these statistics are valid for a wide range of DGPs
- How well can they do? — Take Johansen test as an example:



Introduction

Point (II): Commonly-used Gaussian-based methods cannot use these information

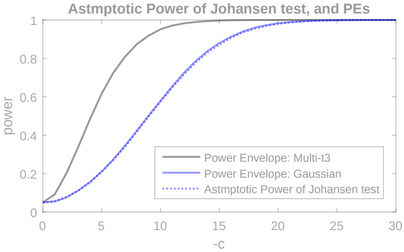
- What are Gaussian-based methods? — Muller (2011)
  - Examples: OLS and t-test for linear regression; the Dickey-Fuller test, the ERS test for unit root; the Johansen test for cointegration; the AR test, the LM test, and the CLR test for weak instruments ...
  - Step I: assume i.i.d ε_t, and f Gaussian $\longrightarrow$ test statistics
  - Step II: show these statistics are valid for a wide range of DGPs
- How well can they do? — Take Johansen test as an example:



Introduction

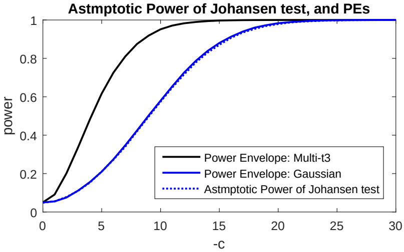
Point (II): Commonly-used Gaussian-based methods cannot use these information

- What are Gaussian-based methods? — Muller (2011)
  - Examples: OLS and t-test for linear regression; the Dickey-Fuller test, the ERS test for unit root; the Johansen test for cointegration; the AR test, the LM test, and the CLR test for weak instruments ...
  - Step I: assume i.i.d ε_t, and f Gaussian $\longrightarrow$ test statistics
  - Step II: show these statistics are valid for a wide range of DGPs
- How well can they do? — Take Johansen test as an example:



Point (II): Commonly-used Gaussian-based methods cannot use these information

- What are Gaussian-based methods? — Muller (2011)
  - Examples: OLS and t-test for linear regression; the Dickey-Fuller test, the ERS test for unit root; the Johansen test for cointegration; the AR test, the LM test, and the CLR test for weak instruments ...
  - Step I: assume i.i.d ε_t, and f Gaussian $\longrightarrow$ test statistics
  - Step II: show these statistics are valid for a wide range of DGPs
- How well can they do? — Take Johansen test as an example:



Introduction

Point (III): The traditional semiparametric approach does not work in nonstandard econometric problems ...

- What is the traditional semiparametric approach?
  - References: Bickel (1982), Newey (1990), Van der Vaart (1991), Andrews (1994), Bickel et al. (1998) ...
  - Step I: Semiparametric efficiency bounds:
  - Step II: Feasible inference based on $\hat{f}$
- What are nonstandard econometric problems?
  - Name comes from Muller & Norets *(forthcoming in Econometrica)*.
  - Feature: nonstandard limit behavior
  - Examples: near unit root, cointegration, weak instrument, predictive regression with a near unit root predictor, moment inequalities ...
- Why does it not work in these nonstandard problems?
  - Tangent space is hard (if not impossible) to calculate
  - Jansson (2008) for unit root testing problem

Introduction

Point (III): The traditional semiparametric approach does not work in nonstandard econometric problems ...

- What is the traditional semiparametric approach?
  - References: Bickel (1982), Newey (1990), Van der Vaart (1991), Andrews (1994), Bickel et al. (1998) ...
  - Step I: Semiparametric efficiency bounds:
  - Step II: Feasible inference based on $\hat{f}$
- What are nonstandard econometric problems?
  - Name comes from Muller & Norets *(forthcoming in Econometrica)*.
  - Feature: nonstandard limit behavior
  - Examples: near unit root, cointegration, weak instrument, predictive regression with a near unit root predictor, moment inequalities ...
- Why does it not work in these nonstandard problems?
  - Tangent space is hard (if not impossible) to calculate
  - Jansson (2008) for unit root testing problem

Introduction

Point (III): The traditional semiparametric approach does not work in nonstandard econometric problems ...

- What is the traditional semiparametric approach?
  - References: Bickel (1982), Newey (1990), Van der Vaart (1991), Andrews (1994), Bickel et al. (1998) ...
  - Step I: Semiparametric efficiency bounds:
  - Step II: Feasible inference based on $\hat{f}$
- What are nonstandard econometric problems?
  - Name comes from Muller & Norets *(forthcoming in Econometrica)*.
  - Feature: nonstandard limit behavior
  - Examples: near unit root, cointegration, weak instrument, predictive regression with a near unit root predictor, moment inequalities ...
- Why does it not work in these nonstandard problems?
  - Tangent space is hard (if not impossible) to calculate
  - Jansson (2008) for unit root testing problem

Introduction

Point (III): The traditional semiparametric approach does not work in nonstandard econometric problems ...

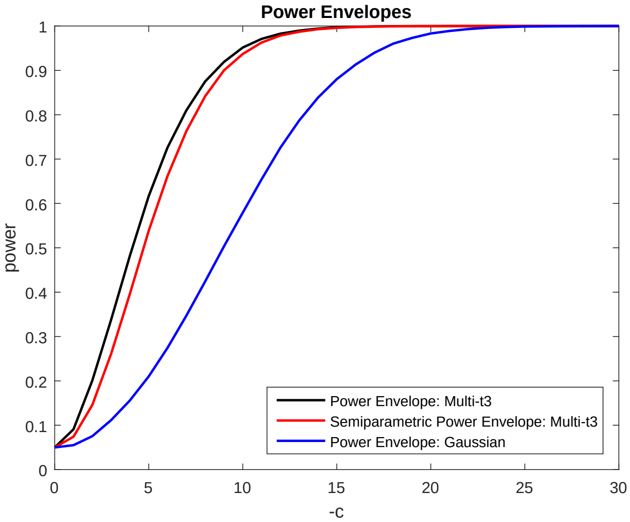
- What is the traditional semiparametric approach?
  - References: Bickel (1982), Newey (1990), Van der Vaart (1991), Andrews (1994), Bickel et al. (1998) ...
  - Step I: Semiparametric efficiency bounds:
  - Step II: Feasible inference based on $\hat{f}$
- What are nonstandard econometric problems?
  - Name comes from Muller & Norets *(forthcoming in Econometrica)*.
  - Feature: nonstandard limit behavior
  - Examples: near unit root, cointegration, weak instrument, predictive regression with a near unit root predictor, moment inequalities ...
- Why does it not work in these nonstandard problems?
  - Tangent space is hard (if not impossible) to calculate
  - Jansson (2008) for unit root testing problem

Main Contributions

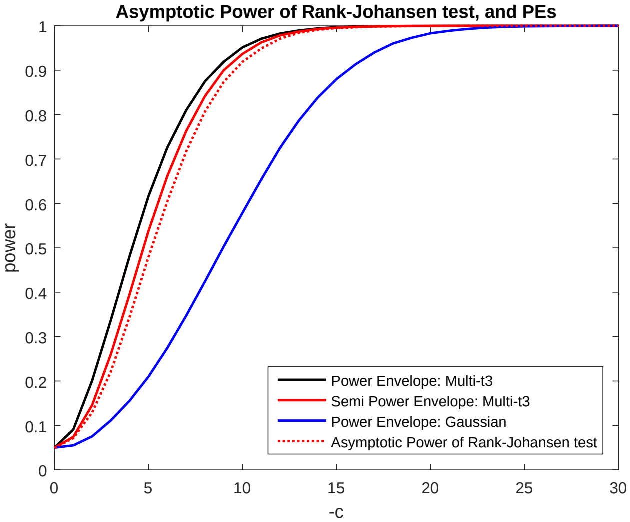
In this paper, I

- Provide a new semiparametric approach
    - to derive the semiparametric bounds (in nonstandard ecnometric problems, the semiparametric power envelopes)
    - leading to results that are linked to two nonparametric literature
        - nonparametrically estimated density $\hat{f}$
        - rank statistic
- Propose a new way to construct statistical inferences based on componentwise rank statistics for multivariate cases without imposing additional assumptions
- Apply them to two problems
    - Cointegration
    - Weak instrument

Main Contributions: e.g., cointegration case



Main Contributions: e.g., cointegration case



Roadmap

(1) The New Approach — Cointegration
(2) Multivariate Rank Statistics — Cointegration
(3) Weak Instrument
(4) Conclusions and Future Research

The New Approach — Cointegration

The New Approach — Step I: Model Setup

- Consider a VAR model expressed in the error correction form

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t, \quad t = 1, ..., T$$

  - Δ denotes first-order differencing
  - y_t is a *two*-dimensional time series with $I(1)$ components
  - ε_t is a *two*-dimensional i.i.d. innovations with joint density f
- Assumptions on f
  - f is absolutely continuous
  - The Fisher-information matrix J_f is finite
  - The covariance matrix Σ is positive definite and finite.
- Denote the rank of $\Pi \in \mathbb{R}^{2\times 2}$ by r, we are interested in testing

$$H_0 : r = 0 \quad \text{against} \quad H_1 : r > 0$$

The New Approach — Step I: Model Setup

- Consider a VAR model expressed in the error correction form

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t, \quad t = 1, ..., T$$

  - Δ denotes first-order differencing
  - y_t is a *two*-dimensional time series with $I(1)$ components
  - ε_t is a *two*-dimensional i.i.d. innovations with joint density f
- Assumptions on f
  - f is absolutely continuous
  - The Fisher-information matrix J_f is finite
  - The covariance matrix Σ is positive definite and finite.
- Denote the rank of $\Pi \in \mathbb{R}^{2\times 2}$ by r, we are interested in testing

$$H_0 : r = 0 \quad \text{against} \quad H_1 : r > 0$$

The New Approach — Step I: Model Setup

- Consider a VAR model expressed in the error correction form

$$\Delta y_t = \Pi y_{t-1} + \varepsilon_t, \quad t = 1, ..., T$$

  - Δ denotes first-order differencing
  - y_t is a *two*-dimensional time series with $I(1)$ components
  - ε_t is a *two*-dimensional i.i.d. innovations with joint density f

- Assumptions on f
  - f is absolutely continuous
  - The Fisher-information matrix J_f is finite
  - The covariance matrix Σ is positive definite and finite.

- Denote the rank of $\Pi \in \mathbb{R}^{2\times 2}$ by r, we are interested in testing

$$H_0 : r = 0 \ \text{ against } \ H_1 : r > 0$$

The New Approach — Step II: Preliminaries

- The nonparametric modeling

$$f_\eta^{(T)}(\varepsilon_t) = f(\varepsilon_t)\left(1 + \frac{1}{\sqrt{T}}\sum_{k=1}^{\infty}\eta_k b_k(\varepsilon_t)\right)$$

  - $e, b_1(e), b_2(e), \ldots$ is an orthonormal basis of space L_2^f
  - $\eta = (\eta_1 \ \eta_2 \ \ldots)'$ is the localized perturbation parameter
- Localization of the parameter of interest

$$\Pi = \frac{C}{T}$$

- We are interested in testing

$$H_0 : C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The New Approach — Step II: Preliminaries

- The nonparametric modeling

$$f_{\eta}^{(T)}(\varepsilon_t) = f(\varepsilon_t)\left(1 + \frac{1}{\sqrt{T}}\sum_{k=1}^{\infty}\eta_k b_k(\varepsilon_t)\right)$$

  - $e, b_1(e), b_2(e), \ldots$ is an orthonormal basis of space L_2^f
  - $\eta = (\eta_1 \ \eta_2 \ \ldots)'$ is the localized perturbation parameter

- Localization of the parameter of interest

$$\Pi = \frac{C}{T}$$

- We are interested in testing

$$H_0 : C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The New Approach — Step II: Preliminaries

- The nonparametric modeling

$$f_{\eta}^{(T)}(\varepsilon_t) = f(\varepsilon_t)\left(1 + \frac{1}{\sqrt{T}}\sum_{k=1}^{\infty}\eta_k b_k(\varepsilon_t)\right)$$

  - $e, b_1(e), b_2(e), \ldots$ is an orthonormal basis of space L_2^f
  - $\eta = (\eta_1 \;\; \eta_2 \;\; \ldots)'$ is the localized perturbation parameter

- Localization of the parameter of interest

$$\Pi = \frac{C}{T}$$

- We are interested in testing

$$H_0 : C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Introduction — Limit Experiment

- What is limit experiment?
  - A sequence of experiments is defined to converge to a limit experiment if the sequence of likelihood ratio processes converges marginally in distribution to the likelihood ratio process of the limit experiment — Van der Vaart (2000)
- Why do we need it?
  - Asymptotic Representation Theorem shows that there is always an asymptotic representative in the limit experiment of your statistical inference in the sequence of experiments (indexed by T) as $T \to \infty$
  - A limit experiment is always statistically easier

- What is limit experiment?
  - A sequence of experiments is defined to converge to a limit experiment if the sequence of likelihood ratio processes converges marginally in distribution to the likelihood ratio process of the limit experiment — Van der Vaart (2000)
- Why do we need it?
  - Asymptotic Representation Theorem shows that there is always an asymptotic representative in the limit experiment of your statistical inference in the sequence of experiments (indexed by T) as $T \to \infty$
  - A limit experiment is always statistically easier

- What is limit experiment?
  - A sequence of experiments is defined to converge to a limit experiment if the sequence of likelihood ratio processes converges marginally in distribution to the likelihood ratio process of the limit experiment — Van der Vaart (2000)
- Why do we need it?
  - Asymptotic Representation Theorem shows that there is always an asymptotic representative in the limit experiment of your statistical inference in the sequence of experiments (indexed by T) as $T \to \infty$
  - A limit experiment is always statistically easier

The New Approach — Step III: Limit Experiment

The limit likelihood ratio is of the Locally Asymptotically Brownian Functional (LABF) form (Jaganathan (1995))

$$\mathcal{L}(C,\eta) = \Delta(C,\eta) - \frac{1}{2}\mathcal{Q}(C,\eta)$$

with

$$\Delta(C,\eta) = \int_0^1 [CW_\varepsilon(u)]' \mathrm{d}W_{\ell_f}(u) + \eta' W_b(1),$$

$$\mathcal{Q}(C,\eta) = \int_0^1 [CW_\varepsilon(u)]' J_f C W_\varepsilon(u)\mathrm{d}u + 2\eta' \int_0^1 J_{bf} C W_\varepsilon(u)\mathrm{d}u + \eta'\eta.$$

- W_ε, W_{ℓ_f}, and W_b are Brownian motions under H_0, with dimensions 2, 2, and ∞
- J_f and J_{fb} are Fisher information matrix

The New Approach — Step IV: Structural Limit Experiment

The structural version of the limit experiment is

$$
\begin{aligned}
\mathrm{d}W_\varepsilon(u) &= C W_\varepsilon(u)\mathrm{d}u + \mathrm{d}Z_\varepsilon(u) \\
\mathrm{d}W_b(u) &= J_{bf} C W_\varepsilon(u)\mathrm{d}u + \eta \mathrm{d}u + \mathrm{d}Z_b(u) \\
\mathrm{d}W_{\ell_f}(u) &= J_f C W_\varepsilon(u)\mathrm{d}u + J_{fb}\eta \mathrm{d}u + \mathrm{d}Z_{\ell_f}(u)
\end{aligned}
$$

- $u \in [0,1]$ denotes the time
- W_ε, W_{ℓ_f}, and W_b are observed processes
- Z_ε, Z_{ℓ_f}, and Z_b are multivariate Brownian motions
- J_f, J_{fb}, and J_{bf} are known matrices
- η are nuisance parameters
- We are interested in testing

$$H_0 : C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

- Define "taking bridge" operator $B^{W(u)} := W(u) - W(1)u$, then

$$B^{W(u)+\zeta u} = W(u) + \zeta u - (W(1) + \zeta)u = W(u) - (W(1))u = B^{W(u)}$$

- $\mathcal{M} = \sigma\left(W_{\varepsilon}(u), B_b(u)\right)$ is invariant
- Moreover

Theorem 1

$\mathcal{M} = \sigma\left(W_{\varepsilon}(u), B_b(u)\right)$ is the Maximal Invariant statistic.

The New Approach — Step VI: Semiparametric Power Envelope

- The log-likelihood ratio of $\mathcal{M}$ is

$$\mathcal{L}_{\mathcal{M}}(C) = \mathrm{E}\left[\mathcal{L}(C,\eta)|\mathcal{M}\right] = \Delta_{\mathcal{M}}(C) - \frac{1}{2}\mathcal{Q}_{\mathcal{M}}(C)$$

with

$$\Delta_{\mathcal{M}}(C) = \int (CW_{\varepsilon})'\mathrm{d}\left[B_{\ell_f} + \Sigma^{-1}W_{\varepsilon}(1)\right]$$

$$\mathcal{Q}_{\mathcal{M}}(C) = \int (CW_{\varepsilon})'J_f CW_{\varepsilon} + \int (CW_{\varepsilon})'\left(\Sigma^{-1} - J_f\right)\int CW_{\varepsilon}$$

- The likelihood test

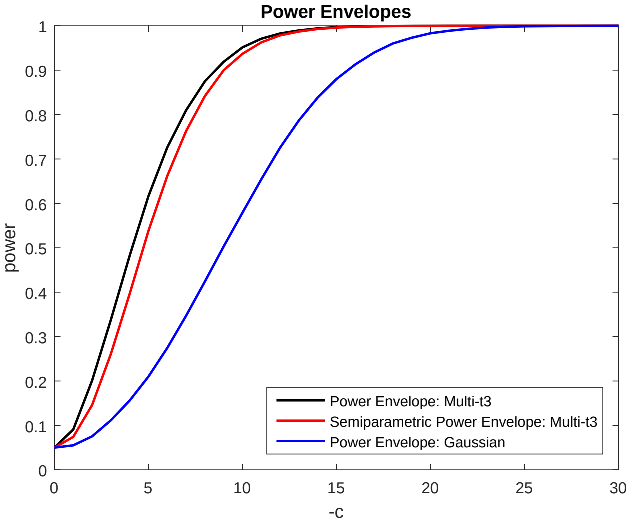
$$\phi(C,\alpha) := \mathbb{1}\left\{\mathcal{L}_{\mathcal{M}}(C) > \kappa_{\alpha}(C)\right\}$$

where $\kappa_{\alpha}(C)$ is the $1-\alpha$ quantile of $\mathcal{L}_{\mathcal{M}}(C)$

- By Neyman-Pearson lemma, the semiparametric power envelope is

$$\Psi(C,\alpha) = \mathrm{E}\left[\phi(C,\alpha)\frac{\mathrm{d}\mathbb{P}_{C,\eta}}{\mathrm{d}\mathbb{P}_{0,0}}\right]$$

Semiparametric Power Envelope Plot



Some Remarks

- Gaussian-based methods are based on $\sigma(W_\varepsilon) \subset \mathcal{M} = \sigma(W_\varepsilon, W_{\ell_f})$
- This approach works for all LAN, LAMN, and LABF models
- The semiparametric bound links with two different literature:
  - Recall the term $\int (CW_\varepsilon)' \mathrm{d}B_{\ell_f} \longrightarrow$ rank-based inference
  - Define $\widetilde{W}_\varepsilon = W_\varepsilon - \int W_\varepsilon$
  - This term equals $\int (C\widetilde{W}_\varepsilon)' \mathrm{d}W_{\ell_f} \longrightarrow \hat{f}$-based inference
- The Structural Limit Experiment is the key of this approach, which can
  - suggest the way to eliminate nuisance parameters
  - suggest the way to construct test statistics (Example: weak instrument)

Multivariate Rank Statistics — Cointegration

Multivariate Rank — Rank-Based Scores

- Standardized error $\nu_t = \Sigma^{-\frac{1}{2}}\varepsilon_t \in \mathbb{R}^p$
- Denote the estimates of ν_t under the null hypothesis by $\hat{\nu}_t$
- Let $R_{i,t}$ be the rank of $\hat{\nu}_{i,t}$ among $\{\hat{\nu}_{i,1}, ..., \hat{\nu}_{i,T}\}$, $i = 1, ..., p$
- Marginal reference density $g_i(\cdot)$ with CDF G_i, inv-CDF G_i^{-1} and score ℓ_{g_i}
- The multivariate rank-based score

$$\ell_g(R_t) = \left(\ell_{g_1}\left[G_1^{-1}\left(\frac{R_{1,t}}{T+1}\right)\right] \cdots \ell_{g_p}\left[G_p^{-1}\left(\frac{R_{p,t}}{T+1}\right)\right]\right)'$$

- Partial sum processes

$$W_\nu^{(T)}(u) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[uT]}\hat{\nu}_t \quad \text{and} \quad B_{\ell_g}^{(T)}(u) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[uT]}\ell_g\left(R_t\right)$$

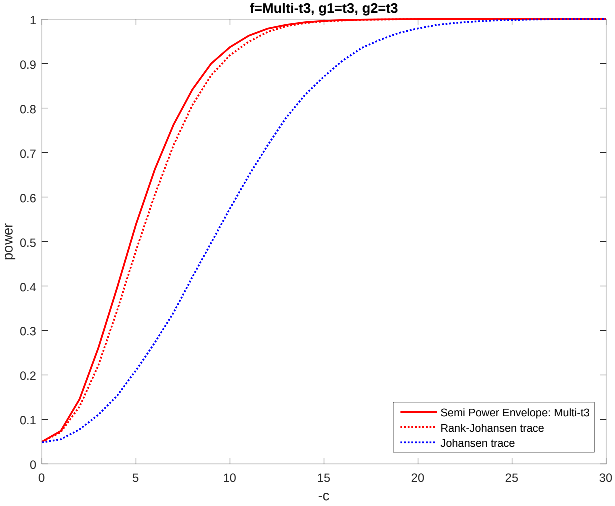
which converges to W_ν and B_{ℓ_g}, respectively.

Multivariate Rank — Two Methods

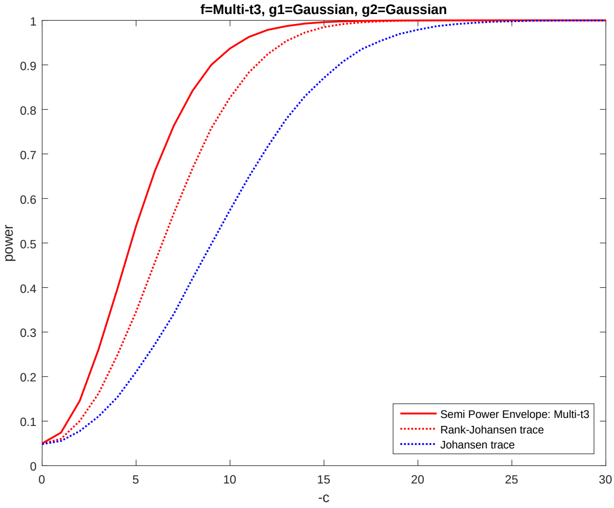
- Method 1 — based on the likelihood of $\sigma(W_\varepsilon, B_{\ell_g})$
  - Pros: attractive theoretical properties, freely choose g_i
  - Cons: hard to implement, hard to study by Monte Carlo
- Method 2 — based on replacement by B_{ℓ_g}
  - Pros: easy to implement, attractive properties in simulations
  - Con: not so freely in choosing g_i
- Implement of method 2
  - replace $W_\varepsilon(u)$ by $\widehat{\Sigma}^{\frac{1}{2}} W_\nu^{(T)}(u)$
  - replace $W_{\ell_f}(u)$ by $\widehat{\Sigma}^{-\frac{1}{2}} B_{\ell_g}^{(T)}(u)$

- Then we have the rank-based version of the Johansen trace test
- Properties
  - Valid (or, asymptotically of size α)
  - Asymptotically invariant
  - Chernoff-Savage result (simulations)
  - Optimal when components of ν_t are independent

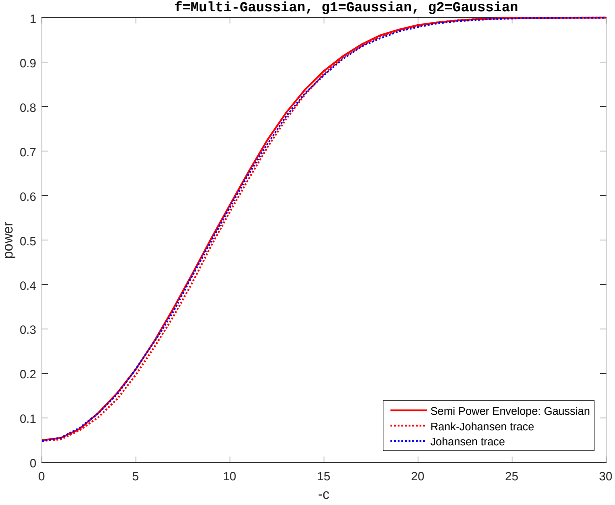
Cointegration — Monte Carlo study with $T = 2500$

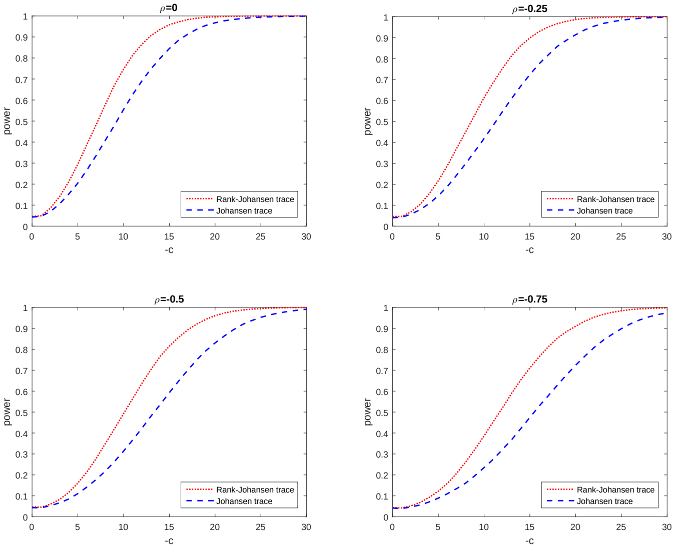


Cointegration — Monte Carlo study with $T = 2500$



Cointegration — Monte Carlo study with $T = 2500$




ρ=0
ρ=-0.25
ρ=-0.5
ρ=-0.75
power
-c
Rank-Johansen trace
Johansen trace

Weak Instrument

Weak Instrument — Model Setup

- The model is

$$y_{1t} = y_{2t}\beta + \varepsilon_{1,t}$$
$$y_{2t} = z_t'\pi + \varepsilon_{2,t}, \quad t = 1, ..., T$$

  - y_{1t} is a scalar dependent variable
  - y_{2t} is a scalar endogenous variable
  - ε_t is a *two*-dimensional i.i.d. innovations with joint density f
  - $z_t \in \mathbb{R}^k$ are instrument variables
  - $\pi \in \mathbb{R}^k$ is a nuisance parameter
  - Denote by ρ the correlation between $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$
- We are interested in testing

$$H_0 : \beta = \beta_0 \ \text{ against } \ H_1 : \beta \neq \beta_0$$

Weak Instrument — Structural Limit Experiment

- Weak IV assumption: $\pi = c/\sqrt{T}$, β is fixed (Staiger and Stock (1997))
- Define $z^{(T)}(u) = z_{[uT]}$
- As $T \to \infty$, $z^{(T)}(u) \to z_u$
- Define $D_{zz} = \int_0^1 z_u z_u' \mathrm{d}u$
- Structural limit experiment, with $a(\beta) = \begin{pmatrix} \beta \\ 1 \end{pmatrix}$

$$\mathrm{d}W_\varepsilon(u) = a(\beta)(c'z_u)\mathrm{d}u + \mathrm{d}Z_\varepsilon(u) \tag{2}$$

$$\mathrm{d}W_{\ell_f}(u) = J_f a(\beta)(c'z_u)\mathrm{d}u + \mathrm{d}Z_{\ell_f}(u) \tag{3}$$

- Gaussian case, where (3)=(2). Write (2) componentwisely

$$\mathrm{d}W_{\varepsilon_1}(u) = \beta c' z_u \mathrm{d}u + \mathrm{d}Z_{\varepsilon_1}(u)$$

$$\mathrm{d}W_{\varepsilon_2}(u) = c' z_u \mathrm{d}u + \mathrm{d}Z_{\varepsilon_2}(u)$$

Statistics should be based on $\mathrm{d}W_{\varepsilon_1} - \beta_0 \mathrm{d}W_{\varepsilon_2}$ and $\hat{c}_{H_0}$

Weak Instrument — Statistics for Gaussian Case

- Statistics based on $\mathrm{d}W_{\varepsilon_1} - \beta_0 \mathrm{d}W_{\varepsilon_2}$ and $\hat{c}_{H_0}$ are
  - $\mathcal{S}_\phi = D_{zz}^{-1/2} \int_0^1 z_u \mathrm{d}\left[W_\varepsilon(u)' b_0\right] \cdot \left(b_0' \Sigma b_0\right)^{-1/2}, \qquad b_0 = \begin{pmatrix} 1 & -\beta_0 \end{pmatrix}'$
  - $\mathcal{T}_\phi = D_{zz}^{-1/2} \int_0^1 z_u \mathrm{d}\left[W_\varepsilon(u)' \Sigma^{-1} a_0\right] \cdot \left(a_0' \Sigma^{-1} a_0\right)^{-1/2} \qquad a_0 = \begin{pmatrix} \beta_0 & 1 \end{pmatrix}'$
- The limit representative of
  - The AR test by Anderson and Rubin (1949) is

$$\mathcal{AR}_\phi = \mathcal{S}_\phi' \mathcal{S}_\phi$$

  - The LM test by Kleibergen (2002) is

$$\mathcal{LM}_\phi = (\mathcal{S}_\phi' \mathcal{T}_\phi)^2 / \mathcal{T}_\phi' \mathcal{T}_\phi$$

  - The CLR test by Moreira (2003) is

$$\mathcal{CLR}_\phi = \frac{1}{2}\left[\mathcal{S}_\phi' \mathcal{S}_\phi - \mathcal{T}_\phi' \mathcal{T}_\phi + \sqrt{(\mathcal{S}_\phi' \mathcal{S}_\phi - \mathcal{T}_\phi' \mathcal{T}_\phi)^2 + 4(\mathcal{S}_\phi' \mathcal{T}_\phi)^2}\,\right]$$

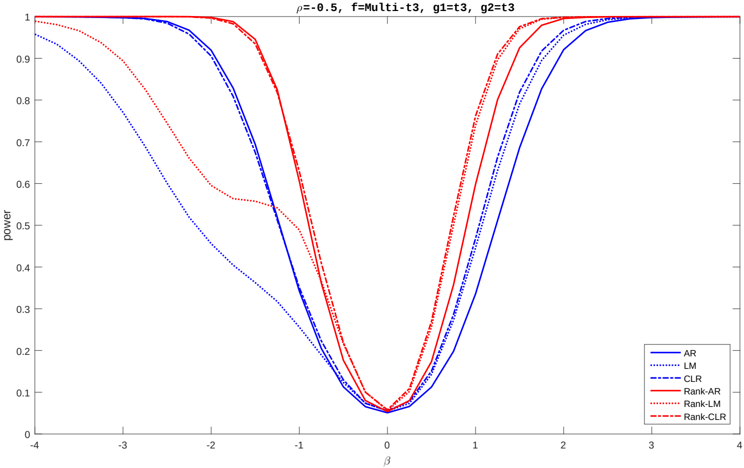
Weak Instrument — Statistics for General Case

- The Structural limit experiment (3)

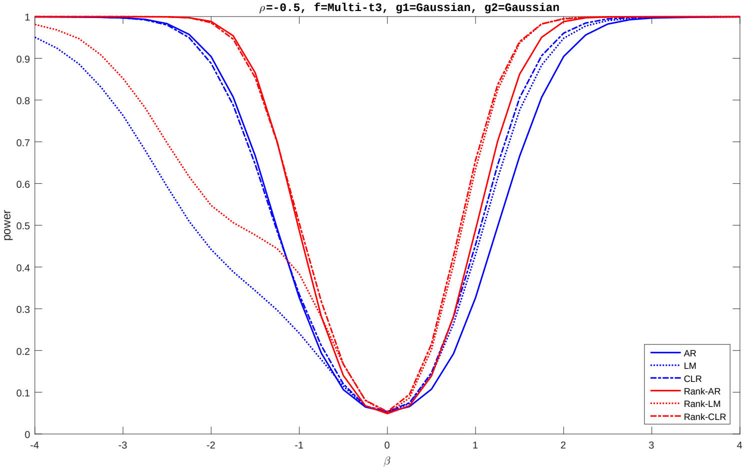
$$J_f^{-1}\begin{bmatrix}\mathrm{d}W_{\ell_f,1}(u)\\ \mathrm{d}W_{\ell_f,1}(u)\end{bmatrix} = \begin{bmatrix}\beta c' z_u\\ c' z_u\end{bmatrix}\mathrm{d}u + J_f^{-1}\begin{bmatrix}\mathrm{d}Z_{\ell_f,1}(u)\\ \mathrm{d}Z_{\ell_f,1}(u)\end{bmatrix}$$

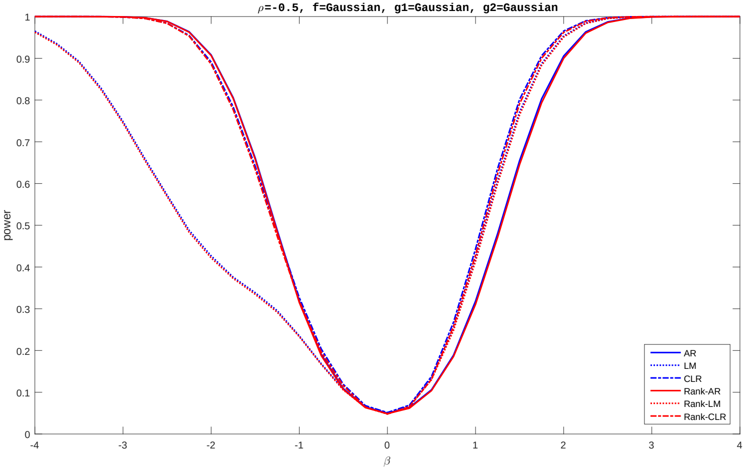
- Similar statistics are
  - $\mathcal{S}_f = D_{zz}^{-1/2}\int_0^1 z_u \mathrm{d}\left[B_{\ell_f}(u)' J_f^{-1} b_0\right]\cdot\left(b_0' J_f^{-1} b_0\right)^{-1/2}$
  - $\mathcal{T}_f = D_{zz}^{-1/2}\int_0^1 z_u \mathrm{d}\left[B_{\ell_f}(u)' a_0\right]\cdot\left(a_0' J_f a_0\right)^{-1/2}$
- The limit representative of
  - The AR test is $\mathcal{AR}_f = \mathcal{S}_f'\mathcal{S}_f$
  - The LM test is $\mathcal{LM}_f = (\mathcal{S}_f'\mathcal{T}_f)^2/\mathcal{T}_f'\mathcal{T}_f$
  - The CLR test is $\mathcal{CLR}_f = 0.5[\mathcal{S}_f'\mathcal{S}_f - \mathcal{T}_f'\mathcal{T}_f + \sqrt{(\mathcal{S}_f'\mathcal{S}_f - \mathcal{T}_f'\mathcal{T}_f)^2 + 4(\mathcal{S}_f'\mathcal{T}_f)^2}]$
- Then, use Method 2 for Multivariate Rank statistics

Weak Instrument — Monte Carlo study with $T = 2500$

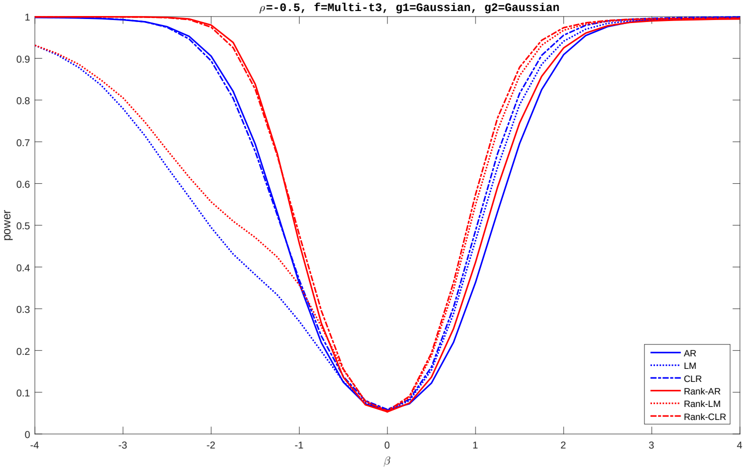


Weak Instrument — Monte Carlo study with $T = 2500$




ρ=-0.5, f=Gaussian, g1=Gaussian, g2=Gaussian
power
β
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4
AR
LM
CLR
Rank-AR
Rank-LM
Rank-CLR

Weak Instrument — Monte Carlo study with $T = 250$



Conclusions and Future Research

- Conclusions
    - A new semiparametric approach
    - A new statistic construction based on multivariate rank statistics
    - Application to cointegration
    - Application to weak instruments
- Future research
    - Other nonstandard econometric problems (e.g. moment inequalities)
    - Other types of semiparametric models
    - Conjecture: little information contained in the copula structure