Semiparametric Cointegrating Rank Selection

Xu Cheng

Peter C. B. Phillips

Department of Economics
Yale University

Workshop on Current Trends and Challenges
in Model Selection and Related Areas
Vienna, July 2008

Papers and Outline

- Cheng & Phillips (2008a)
  "Semiparametric cointegrating rank selection"
  - consistent cointegrating rank estimation by information criteria
  - asymptotics for weakly dependent innovations
  - simulation
- Cheng & Phillips (2008b)
  "Cointegrating rank selection in models with time varying variance"
  - robust to unconditional heterogeneity of unknown form
  - asymptotics under time varying variances
  - empirical application on exchange rate dynamics and simulation

Papers and Outline

- Cheng & Phillips (2008a)
  "Semiparametric cointegrating rank selection"
  - consistent cointegrating rank estimation by information criteria
  - asymptotics for weakly dependent innovations
  - simulation
- Cheng & Phillips (2008b)
  "Cointegrating rank selection in models with time varying variance"
  - robust to unconditional heterogeneity of unknown form
  - asymptotics under time varying variances
  - empirical application on exchange rate dynamics and simulation

Papers and Outline

- Cheng & Phillips (2008a)
  "Semiparametric cointegrating rank selection"
  - consistent cointegrating rank estimation by information criteria
  - asymptotics for weakly dependent innovations
  - simulation
- Cheng & Phillips (2008b)
  "Cointegrating rank selection in models with time varying variance"
  - robust to unconditional heterogeneity of unknown form
  - asymptotics under time varying variances
  - empirical application on exchange rate dynamics and simulation

SP ECM Model

- semiparametric ECM

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

  α and β are $m \times r$ full rank matrices

  - u_t is weakly dependent with mean zero
- general short memory component u_t
  - no specification of VAR lags as in $\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta X_{t-i} + u_t$
  - no specification of the distribution of u_t
  - allow for unconditional unknown heterogeneity in u_t
- permit near integration as well as strict unit roots

SP ECM Model

- semiparametric ECM

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

  α and β are $m \times r$ full rank matrices

  – u_t is weakly dependent with mean zero

- general short memory component u_t
  – no specification of VAR lags as in $\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta X_{t-i} + u_t$
  – no specification of the distribution of u_t
  – allow for unconditional unknown heterogeneity in u_t

- permit near integration as well as strict unit roots

Cointegrating rank determination

- fitted model

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

- information criteria

$$IC(r) = \log\left|\widehat{\Sigma}\left(r\right)\right| + C_n n^{-1}\left(2mr - r^2\right), \quad 0 \le r \le m$$

  - $\widehat{\Sigma}\left(r\right)$ is the residual covariance matrix from reduced rank regression
  - penalty C_n: 2 (AIC), $\log\left(n\right)$ (BIC), $c\log\log\left(n\right)$ (HQ)
  - degrees of freedom: $2mr - r^2$

- $\widehat{r} = \arg\min_{0\le r\le m} IC(r)$

Cointegrating rank determination

- fitted model

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

- information criteria

$$IC(r) = \log \left|\widehat{\Sigma}\,(r)\right| + C_n n^{-1}\left(2mr - r^2\right), \qquad 0 \leq r \leq m$$

  - $\widehat{\Sigma}\,(r)$ is the residual covariance matrix from reduced rank regression
  - penalty C_n: 2 (AIC), $\log(n)$ (BIC), $c \log\log(n)$ (HQ)
  - degrees of freedom: $2mr - r^2$

- $\widehat{r} = \arg\min_{0 \leq r \leq m} IC(r)$

Cointegrating rank determination

- fitted model

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

- information criteria

$$IC(r) = \log|\widehat{\Sigma}(r)| + C_n n^{-1}\left(2mr - r^2\right), \quad 0 \le r \le m$$

  - $\widehat{\Sigma}(r)$ is the residual covariance matrix from reduced rank regression
  - penalty C_n: 2 (AIC), $\log(n)$ (BIC), $c\log\log(n)$ (HQ)
  - degrees of freedom: $2mr - r^2$

- $\widehat{r} = \arg\min_{0\le r\le m} IC(r)$

Cointegrating rank determination

- fitted model

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

- information criteria

$$IC(r) = \log|\widehat{\Sigma}(r)| + C_n n^{-1}\left(2mr - r^2\right), \quad 0 \leq r \leq m$$

  - $\widehat{\Sigma}(r)$ is the residual covariance matrix from reduced rank regression
  - penalty C_n: 2 (AIC), $\log(n)$ (BIC), $c\log\log(n)$ (HQ)
  - degrees of freedom: $2mr - r^2$

- $\widehat{r} = \arg\min_{0\leq r\leq m} IC(r)$

Outline of Basic Results

$$\begin{aligned} \Delta X_t &= \alpha\beta' X_{t-1} + u_t \\ IC(r) &= \log|\widehat{\Sigma}(r)| + C_n n^{-1}\left(2mr - r^2\right) \end{aligned}$$

- $IC(r)$ is weakly consistent provided $C_n \to \infty$ and $C_n/n \to 0$
- AIC inconsistent, limit distribution given

Literature — Order Estimation

- Semiparametric approaches
  - Phillips (2008) "Unit root model selection"
- Parametric approaches & joint order estimation
  - Johansen (1988, 1991)
  - Phillips and Ploberger (1996), Phillips (1996), Chao and Phillips (1999)
  - Phillips & McFarland (1997)
- Order selection & nonstationarity
  - Tsay (1984), Potscher (1989), Wei (1992), Nielsen (2006), Kapetanios (2004), Wang & Bessler (2005), Poskitt (2000), Harris and Poskitt (2004)

Literature — Time-varying Variance

- literature
  - classical unit root testing

    Hamori and Tokihisa, 1997; Kim *et al*, 2002; Cavaliere, 2004; Cavaliere and Taylor, 2007; Beare, 2007
  - autoregressive models

    Phillips and Xu, 2006; Xu and Phillips, 2008
- SP model choice method
  - robust to time-varying variance
  - no change in implementation

Contribution

$$\text{SP ECM:} \qquad \Delta X_t = \alpha\beta' X_{t-1} + u_t$$

$\mathbf{u}_t$	**standard method**	**our method**
stationary	valid specify lag length	valid avoid misspecification easy to implement
time-varying var	*invalid*	valid same implementation

Reduced Rank Regression

$$\Delta X_t = \alpha\beta' X_{t-1} + u_t$$

- suppose the cointegrating rank is r
- for given β
  - $\widehat{\alpha}(\beta) = S_{01}\beta(\beta' S_{11}\beta)^{-1}$
  - $\widehat{\Sigma}(\beta) = S_{00} - S_{01}\beta\left(\beta' S_{11}\beta\right)^{-1}\beta' S_{10}$
  - notation

$$S_{00} = n^{-1}\sum_{t=1}^{n}\Delta X_t \Delta X_t', \qquad S_{11} = n^{-1}\sum_{t=1}^{n} X_{t-1}X_{t-1}',$$
$$S_{10} = n^{-1}\sum_{t=1}^{n}\Delta X_t X_{t-1}', \qquad S_{10} = S_{01}'$$

- $\widehat{\beta} = \arg\min_{\beta}|\widehat{\Sigma}(\beta)|$, subject to $\widehat{\beta}' S_{11}\widehat{\beta} = I_r$

  $\widehat{\alpha}(\beta) = \widehat{\alpha}(\widehat{\beta})$ and $\widehat{\Sigma}(r) = \widehat{\Sigma}(\widehat{\beta})$

RRR Estimation – as if model correctly specified

- Johansen (1988,1995)
- determinantal equation $\left|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = 0$
  - ordered eigenvalues $1 > \widehat{\lambda}_1 > \cdots > \widehat{\lambda}_m > 0$
  - corresponding eigenvectors $\widehat{V} = [\widehat{v}_1, \cdots, \widehat{v}_m]$, normalized by $\widehat{V}' S_{11} \widehat{V} = I_m$
- $\widehat{\beta} = [\widehat{v}_1, \cdots, \widehat{v}_r]$ and $|\widehat{\Sigma}(r)| = |S_{00}| \Pi_{i=1}^{r} (1 - \widehat{\lambda}_i)$

Information criteria components

- criterion has the form
  - $IC(r) = \log(|S_{00}| \Pi_{i=1}^{r}(1 - \widehat{\lambda}_i)) + C_n n^{-1}(2mr - r^2)$
  - $\widehat{\lambda}_i$ are ordered solutions of $\left|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = 0$
- find SP limits of $\widehat{\lambda}_{i,}$ for $i = 1, ..., m$, using limit theory for S_{11}, S_{10}, S_{00}

Heuristics

- $|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$
- when X_t is stationary, i.e. $r = m$
  - S_{11}, S_{10}, and S_{00} are all $O_p(1) \Rightarrow 0 < \lambda_i < 1$ for all i
- when X_t is full rank integrated, i.e. $r = 0$
  - $S_{11} = O_p(n) \Rightarrow \lambda_i$ decreases to 0 at rate n^{-1} for all i

Heuristics (cont.)

- when $0 < r < m$
  - $\beta' X_t$ is stationary $\Rightarrow 0 < \lambda_i < 1$ for all $1 \leq i \leq r$
  - $\beta'_\perp X_t$ is an $m - r$ vector of unit root time series

    $\Rightarrow \lambda_i$ decreases to 0 at rate n^{-1} for $r + 1 \leq i \leq m$

- same asymptotic orders apply when u_t is
  - weakly dependent
  - with time varying variance

- KEY: for weak consistency of $IC(r)$, only asymptotic order matters

Heuristics (cont.)

- when $0 < r < m$
  - $\beta' X_t$ is stationary $\Rightarrow 0 < \lambda_i < 1$ for all $1 \leq i \leq r$
  - $\beta'_{\perp} X_t$ is an $m - r$ vector of unit root time series

    $\Rightarrow \lambda_i$ decreases to 0 at rate n^{-1} for $r + 1 \leq i \leq m$

- same asymptotic orders apply when u_t is
  - weakly dependent
  - with time varying variance

- KEY: for weak consistency of $IC(r)$, only asymptotic order matters

Heuristics (cont.)

- when $0 < r < m$
  - $\beta' X_t$ is stationary $\Rightarrow 0 < \lambda_i < 1$ for all $1 \leq i \leq r$
  - $\beta'_\perp X_t$ is an $m - r$ vector of unit root time series

    $\Rightarrow \lambda_i$ decreases to 0 at rate n^{-1} for $r + 1 \leq i \leq m$

- same asymptotic orders apply when u_t is
  - weakly dependent
  - with time varying variance

- KEY: for weak consistency of $IC(r)$, only asymptotic order matters

Information Criteria

- recall $IC(r) = \log(|S_{00}| \, \Pi_{i=1}^{r}(1 - \widehat{\lambda}_i)) + C_n n^{-1}(2mr - r^2)$
- we want $IC_{r_0}(r) - IC_{r_0}(r_0) > 0$ for any $r \neq r_0$
- if $r < r_0$
  - $IC_{r_0}(r) - IC_{r_0}(r_0) =$
    $$\underbrace{-\sum_{i=r+1}^{r_0} \log(1 - \hat{\lambda}_i)}_{+\text{ve}} + \underbrace{C_n n^{-1}(r - r_0)(2m - r - r_0)}_{-\text{ve}}$$
  - $IC_{r_0}(r) - IC_{r_0}(r_0) > 0$ requires $C_n n^{-1} \to 0$ and poor fit dominates

Information Criteria (cont.)

- recall $IC(r) = \log(|S_{00}|\,\Pi_{i=1}^{r}(1-\widehat{\lambda}_i)) + C_n n^{-1}(2mr - r^2)$
- we want $IC_{r_0}(r) - IC_{r_0}(r_0) > 0$ for any $r \neq r_0$
- if $r > r_0$
  - $IC_{r_0}(r) - IC_{r_0}(r_0) =$
    $$\underbrace{\sum_{i=r_0+1}^{r} \log(1-\hat{\lambda}_i)}_{-\text{ve}, \quad O_p(n^{-1})} + \underbrace{C_n n^{-1}(r-r_0)(2m-r-r_0)}_{+\text{ve}}$$
  - $IC_{r_0}(r) - IC_{r_0}(r_0) > 0$ requires $C_n \to \infty$ and penalty dominates

Consistent Information Criteria

Theorem

$IC(r)$ is weakly consistent for selecting the rank of cointegration provided $C_n \to \infty$ and $C_n/n \to 0$.

- BIC and HQ are weakly consistent, but AIC is not
- AIC limit theory
  - no tendency to underestimate
  - tendency to overestimate - just as in lag order

Consistent Information Criteria

Theorem

$IC(r)$ is weakly consistent for selecting the rank of cointegration provided $C_n \to \infty$ and $C_n/n \to 0$.

- BIC and HQ are weakly consistent, but AIC is not
- AIC limit theory
  - no tendency to underestimate
  - tendency to overestimate - just as in lag order

Simulation

- $m = 2$
- $u_t = Au_{t-1} + \varepsilon_t$, $u_t = \varepsilon_t + B\varepsilon_{t-1}$, $u_t = Au_{t-1} + \varepsilon_t + B\varepsilon_{t-1}$
  - $A = \psi I_m$, $B = \phi I_m$
  - $\psi = \phi = 0.4$
- $\varepsilon_t = g(\frac{t}{n})e_t$ and $e_t = iid\ N\left(0, \Sigma_\varepsilon\right)$
  - $\Sigma_\varepsilon = diag\{1 + \theta, 1 - \theta\}$, and $\theta = .25$

Simulation

- choice of variance

1. $g^2(r) = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) 1_{\{r \geq \tau\}}, \quad r \in [0,1],$
2. $g^2(r) = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) 1_{\{\tau \leq r < 1-\tau\}}, \quad r \in [0,1], \ \tau \in [0,1/2],$
3. $g^2(r) = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) r^m, \quad r \in [0,1].$

- In model 1, the break date τ takes values within the set $\{0.1, 0.5, 0.9\}$
- In model 2, τ takes value from $\{0.1, 0.4\}$
- In model 3, we allow for both linear trend and quadratic trend by setting $m \in \{1, 2\}$
- $\delta = \sigma_1/\sigma_0 \in \{0.2, 5\}$

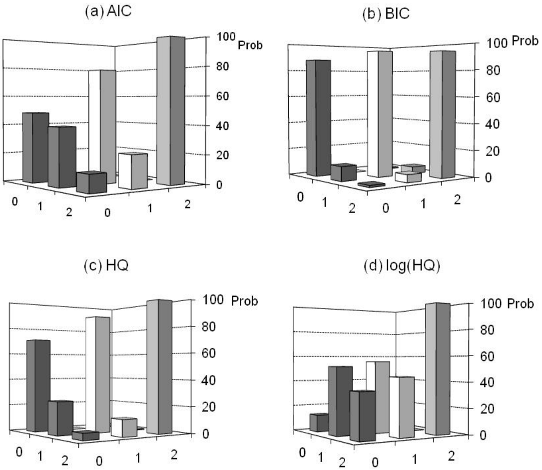

Figure: u_t is AR(1), $\sigma_1 = \sigma_0$, and $n = 100$.

Table 1. u_t follows an AR(1), $g(r) = \sigma_0^2 + \left(\sigma_1^2 - \sigma_0^2\right) 1_{\{\tau \le r < 1-\tau\}}$

			$n = 400$					
			$r_0 = 0$		$r_0 = 1$		$r_0 = 2$	
τ	δ	$\widehat{r}$	AIC	BIC	AIC	BIC	AIC	BIC
0.1	0.2	0	0.50	**0.93**	0.00	0.00	0.00	0.00
		1	0.38	0.06	0.74	**0.96**	0.00	0.00
		2	0.11	0.01	0.26	0.04	1.00	**1.00**
	5	0	0.24	**0.65**	0.00	0.00	0.00	0.00
		1	0.64	0.34	0.75	**0.90**	0.00	0.00
		2	0.12	0.01	0.25	0.10	1.00	**1.00**

Table 2. u_t follows an AR(1), $g(r) = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2)\, 1_{\{\tau \le r < 1-\tau\}}$

			$n = 400$					
			$r_0 = 0$		$r_0 = 1$		$r_0 = 2$	
τ	δ	$\widehat{r}$	AIC	BIC	AIC	BIC	AIC	BIC
0.4	0.2	0	0.38	**0.84**	0.00	0.00	0.00	0.00
		1	0.44	0.14	0.65	**0.92**	0.00	0.00
		2	0.18	0.02	0.35	0.08	1.00	**1.00**
	5	0	0.33	**0.82**	0.00	0.00	0.00	0.00
		1	0.57	0.17	0.75	**0.93**	0.00	0.00
		2	0.09	0.01	0.25	0.07	1.00	**1.00**

Main results

- semiparametric cointegrating rank selection $\Delta X_t = \alpha\beta' X_{t-1} + u_t$
- $\widehat{r} = \arg\min_{0 \leq r \leq m} \left\{ IC(r) = \log \left| \widehat{\Sigma}(r) \right| + C_n n^{-1} \left(2mr - r^2 \right) \right\}$ is weakly consistent provided
  - $C_n \to \infty$
  - $C_n / n \to 0$
- method is robust to persistent heterogeneity and near integration
- easy to implement in practical work