Generalized Local-to-Unity Models

Liyu Dou and Ulrich K. Müller
CUHK Shenzen and Princeton University

July 2020

Motivation

- Many macroeconomic and financial series are highly persistent

  GDP and its aggregates, unemployment, interest rates, price-dividends ratio, etc.

- I(1) model statistically reasonable benchmark

  Unit root tests often inconclusive or only weak rejections

- I(1) persistence leads to alternative (frequentist) econometrics

  Cointegration framework: Granger (1981, 1983), Engle and Granger (1987), Johansen (1988, 1991), Phillips (1988, 1990), Stock and Watson (1993), etc

Fragility of I(1) Inference

- I(1) inference valid or at least conservative for all strongly persistent series?

  No! Elliott (1998): Local-to-unity (LTU) common stochastic trend x_t

  $$(1 - \rho_T L)x_t = u_t, \quad \rho_T = 1 - c/T$$

  invalidates cointegration inference

- Robust inference under local-to-unity asymptotics

  Elliott and Stock (1994), Stock and Watson (1996), Moon and Phillips (2000), Campbell and Yogo (2006), Jansson and Moreira (2006), Mikusheva (2007, 2012), Müller (2014), Moon and Velasco (2014), etc.

Properties of LTU Model

- Asymptotics on unit interval

$$T^{-1/2}(x_{\lfloor \cdot T \rfloor} - x_1) \Rightarrow J_1(\cdot) - J_1(0)$$

where J_1 is continuous time Gaussian AR(1) ("Ornstein Uhlenbeck process") satisfying

$$dJ_1(s) = -cJ_1(s)ds + dW(s)$$

so x_t has AR(1) "long-run" dynamics $\text{Corr}(x_{\lfloor sT \rfloor}, x_{\lfloor rT \rfloor}) \to e^{-c|r-s|}$

- Key parameter c cannot be consistently estimated, $c = 0$ corresponds to I(1) model
  - Arguably good thing, since I(1) model is reasonable benchmark
    Measure of $J_1(\cdot) - J_1(0)$ is equivalent to Wiener measure
  - Leads to more complicated inference

This Paper: Generalized LTU Model

- Fragility of LTU model inference? Generalize to GLTU(p) model with richer long-run dynamics

$$(1-\rho_{T,1}L)\cdots(1-\rho_{T,p}L)x_t = (1-\gamma_{T,1}L)\cdots(1-\gamma_{T,p-1}L)u_t,$$

where $\rho_{T,j} = 1 - c_j/T$ and $\gamma_{T,j} = 1 - g_j/T$, so GLTU(1) $\triangleq$ LTU

- **Theorem 1**

$$T^{-1/2}x_{\lfloor \cdot T \rfloor} \Rightarrow J_p(\cdot)$$

where J_p is stationary Gaussian CARMA($p, p-1$) process with parameters $\{c_j\}_{j=1}^{p}$ and $\{g_j\}_{j=1}^{p-1}$

- Choice of $(p, p-1)$ orders ensures that measure of $J_p(\cdot) - J_p(0)$ remains equivalent to $W(\cdot)$ (while CARMA($p, p-2$), for instance, has measure equivalent to an integrated Wiener process)

This Paper II: Richness of GLTU Model Class

- Fragility of GLTU model inference?

- **Theorem 2**

  - Let G be a given stationary Gaussian continuous time process whose measure of $G(\cdot) - G(0)$ is equivalent to Wiener measure (+ weak technical condition on its spectral density).

  - For any $\varepsilon > 0$, there exists a Gaussian CARMA($p_\varepsilon, p_\varepsilon - 1$) process $J_{p_\varepsilon}(\cdot)$ such that total variation distance between $G(\cdot)$ and $J_{p_\varepsilon}(\cdot)$ is smaller than ε.

Roughly speaking, GLTU class can approximate all stationary forms of persistence that cannot be perfectly discriminated from I(1) model in large samples

This Paper III: Estimation of GLTU(p) Model

- For given fixed N, by continuous mapping theorem

$$\{T^{-1/2}x_{\lfloor jT/N \rfloor}\}_{j=1}^{N} \Rightarrow \{J_p(j/N)\}_{j=1}^{N}.$$

  Asymptotically justified to treat $x_{\lfloor jT/N \rfloor}$ as discretely sampled observations from Gaussian CARMA($p, p-1$) process

- Literature contains suggestions for likelihood evaluation of discretely observed CARMA process

  But involve matrix exponentials and/or complex valued state-space systems

- **Result**: Arbitrarily good approximation of this *limited information* likelihood via straightforward real Kalman filter

  Still only limited information about $\{c_j\}_{j=1}^{p}$ and $\{g_j\}_{j=1}^{p-1}$, so frequentist inference generically hard

Outline of Talk

1. Introduction
2. GLTU Asymptotics
3. Richness Theorem
4. Limited Information Likelihood Approximation
5. Applications

GLTU(p) Model

- Stationary observed series $x_t, t = 1, \ldots, T$ satisfies

$$(1 - \rho_{T,1}L) \cdots (1 - \rho_{T,p}L)x_t = (1 - \gamma_{T,1}L) \cdots (1 - \gamma_{T,p-1}L)u_t,$$

where $\rho_{T,j} = 1 - c_j/T$ and $\gamma_{T,j} = 1 - g_j/T$

  – u_t stationary and I(0) in the sense of $T^{-1/2}\sum_{t=1}^{\lfloor \cdot T \rfloor} u_t \Rightarrow W(\cdot)$

  – Parameter space for $\{c_j\}_{j=1}^{p}$ and $\{g_j\}_{j=1}^{p-1}$: Polynomials

$$\begin{aligned} a(z) &= \textstyle\prod_{j=1}^{p}(c_j + z) = z^p + \sum_{j=1}^{p} a_j z^{p-j} \\ b(z) &= \prod_{j=1}^{p-1}(g_j + z) = z^{p-1} + \sum_{j=0}^{p-2} b_j z^j \end{aligned}$$

have real coefficients a_j, b_j

CARMA($p, p-1$) Model $J_p(\cdot)$

$J_p(\cdot)$ is process on unit interval satisfying

$$J_p(s) = \mathbf{b}'\mathbf{X}(s)$$

where $p \times 1$ process $\mathbf{X}(\cdot)$ satisfies

$$\mathbf{X}(s) = \int_{-\infty}^{s} e^{\mathbf{A}(s-r)}\mathbf{e}dW(r)$$

with

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ 1 \end{pmatrix}, \mathbf{e} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

GLTU(p) Asymptotics

- By standard state space representation of ARMA($p, p-1$) process

$$\begin{aligned} x_t &= \boldsymbol{\theta}_T' \mathbf{V}_t \\ \mathbf{V}_t &= \boldsymbol{\Phi}_T \mathbf{V}_{t-1} + \mathbf{e} u_t \end{aligned}$$

where $\mathbf{V}_t = (v_{t-p+1}, \ldots, v_{t-1}, v_t)'$,

$$\boldsymbol{\Phi}_T = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\phi_{T,p} & -\phi_{T,p-1} & \cdots & -\phi_{T,1} \end{pmatrix}, \ \boldsymbol{\theta}_T = \begin{pmatrix} \theta_{T,0} \\ \theta_{T,1} \\ \vdots \\ \theta_{T,p-2} \\ 1 \end{pmatrix}, \ \mathbf{e} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$\prod_{j=1}^{p} (z - \rho_{T,j}) = z^p + \sum_{j=1}^{p} \phi_{T,j} z^{p-j} \text{ and } \prod_{j=1}^{p-1} (z - \gamma_{T,j}) = z^{p-1} + \sum_{j=0}^{p-2} \theta_{T,j} z^j$$

- But not helpful for GLTU asymptotics, since no obvious convergence in $T^{-1/2} \mathbf{V}_{\lfloor sT \rfloor} = T^{-1/2} \sum_{t=-\infty}^{\lfloor sT \rfloor} \boldsymbol{\Phi}_T^{\lfloor sT \rfloor - t} \mathbf{e} u_t$ for $p > 1$...

GLTU(p) Asymptotics II

- Key insight: Above system can be rewritten as

$$\begin{aligned} x_t &= \mathbf{b}'\mathbf{Z}_t \\ \mathbf{Z}_t &= (\mathbf{I}_p + \mathbf{A}/T)\mathbf{Z}_{t-1} + \mathbf{e}u_t \end{aligned}$$

where $\mathbf{Z}_t \in \mathbb{R}^p$, with $\mathbf{A}$ and $\mathbf{b}$ as in definition of CARMA$(p, p-1)$ process

- Now

$$T^{-1/2}\mathbf{Z}_{\lfloor sT \rfloor} = T^{-1/2} \sum_{t=-\infty}^{\lfloor sT \rfloor} (\mathbf{I}_p + \mathbf{A}/T)^{\lfloor sT \rfloor - t} \mathbf{e}u_t,$$

strongly suggesting $T^{-1/2}\mathbf{Z}_{\lfloor sT \rfloor} \Rightarrow \mathbf{X}(s) = \int_{-\infty}^{s} e^{\mathbf{A}(s-r)}\mathbf{e}dW(r)$

Theorem 1 *GLTU(p) model satisfies* $T^{-1/2}x_{\lfloor \cdot T \rfloor} \Rightarrow J_p(\cdot)$.

Richness of GLTU Model Class

Theorem 2 *Let G be a mean-zero continuous time stationary Gaussian process on the unit interval satisfying*

(i) $G(\cdot) - G(0)$ is absolutely continuous with respect to the measure of W;

(ii) G has a spectral density $f_G : \mathbb{R} \to [0, \infty)$ satisfying $\sup_\lambda (1+\lambda^2) f_G(\lambda) < \infty$ and $\inf_\lambda (1+\lambda^2) f_G(\lambda) > 0$.

For any $\varepsilon > 0$, there exists a CARMA($p_\varepsilon, p_\varepsilon - 1$) process J_{p_ε} such that the total variation distance between the measures of G and J_{p_ε} is smaller than ε.

Proof of Theorem 2

- Spectral density of $J_p(\cdot)$ is $f_p : \mathbb{R} \mapsto \mathbb{R}$

$$f_p(\lambda) = \frac{\omega^2}{2\pi} \frac{|b(i\lambda)|^2}{|a(i\lambda)|^2} = \frac{\omega^2}{2\pi} \frac{\prod_{j=1}^{q-1}(\lambda^2 + g_i^2)}{\prod_{j=1}^{q}(\lambda^2 + c_i^2)}$$

  Can approximate spectral density f_G of G over compact intervals arbitrarily well

- Tail of spectral density governs high frequency properties

  Leverage classic results of Ibragimov and Rozanov (1978) on equivalence (but not approximability) of Gaussian processes

Limited Information Likelihood Approximation

- **Corollary of Theorem 1** *For fixed* N,

$$\{T^{-1/2}x_{\lfloor jT/N \rfloor}\}_{j=1}^{N} \Rightarrow \{J_p(j/N)\}_{j=1}^{N}.$$

- For $t = 1, \ldots, T_0$ and some large T_0, define Gaussian ARMA$(p, p-1)$ process

$$(1-\rho_{T_0,1}L)\cdots(1-\rho_{T_0,p}L)x_t^0 = (1-\gamma_{T_0,1}L)\cdots(1-\gamma_{T_0,p-1}L)u_t^0$$

with $\rho_{T_0,j} = 1 - c_j/T_0$ and $\gamma_{T_0,j} = 1 - g_j/T_0$.

Also satisfies $\{T_0^{-1/2}x^0_{\lfloor jT_0/N \rfloor}\}_{j=1}^{N} \Rightarrow \{J_p(j/N)\}_{j=1}^{N}$.

- Write x_t^0 in standard state space form, treat all observations other than $\{x^0_{\lfloor jT_0/N \rfloor}\}_{j=1}^{N}$ as missing, and employ Kalman filter

Approximates joint Gaussian likelihood of $\{J_p(j/N)\}_{j=1}^{N}$ arbitrarily well as $T_0 \to \infty$

Two Applications

1. Fragility of LTU inference of Campbell and Yogo (2006)

2. Bayesian limited-information inference about PPP deviations in GLTU(p) model

Campbell-Yogo (2006) in GLTU(2) Model

- Model for stock returns y_t and price dividend ratio x_t

$$\begin{aligned} y_t &= \mu_y + \beta x_{t-1} + e_t, \\ (1 - \rho_T L)(x_t - \mu) &= u_t \end{aligned}$$

where (e_t, u_t) are weakly dependent with long-run correlation r_{eu}.

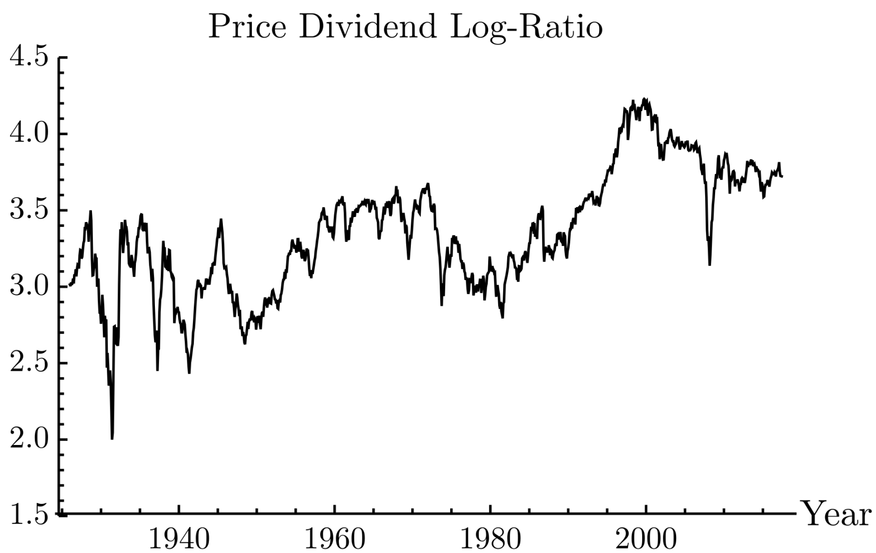
- Suppose instead

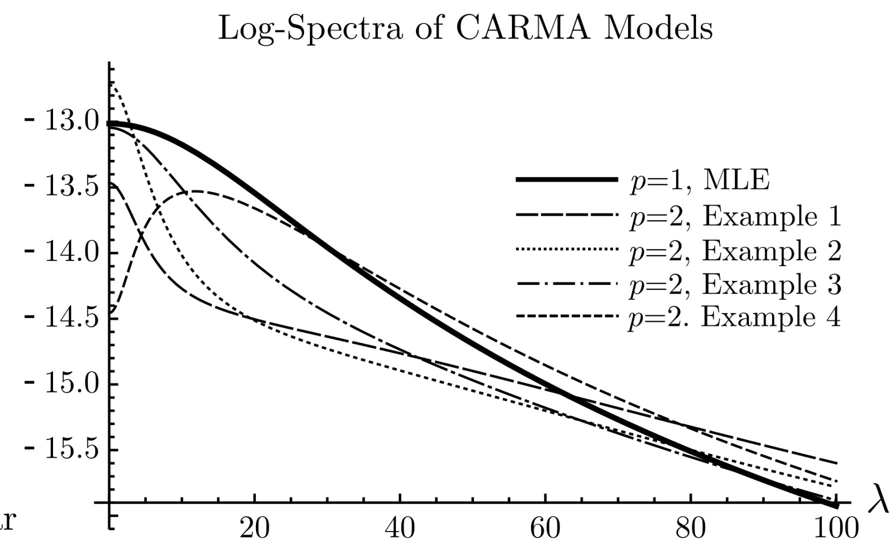
$$(1 - \rho_{T,1} L)(1 - \rho_{T,2} L)(x_t - \mu) = (1 - \gamma_{T,1} L) u_t$$

governed by parameters (c_1, c_2, g_1) instead of single LTU parameter c.

- Does Campbell-Yogo test still yield valid inference? If not, how badly is its size distorted?

CRSP Log Price Dividend Ratio 1926:1-2018:6





⇒ Obtain set of "empirically plausible" values for (c_1, c_2, g_1) that are within 2 limited information log-likelihood points of MLE, $N = 50$

Campbell-Yogo (2006) in GLTU(2) Model

Example No.	1	2	3	4
Value of c_1	70.9	70.0	60.6	29.8
Value of c_2	4.4	4.4	11.1	6.0
Value of g_1	7.3	11.7	24.3	3.3
Null rejection probability	**49.8%**	**46.6%**	**41.2%**	**40.3%**

Severely invalid inference under empirically plausible departures of LTU model

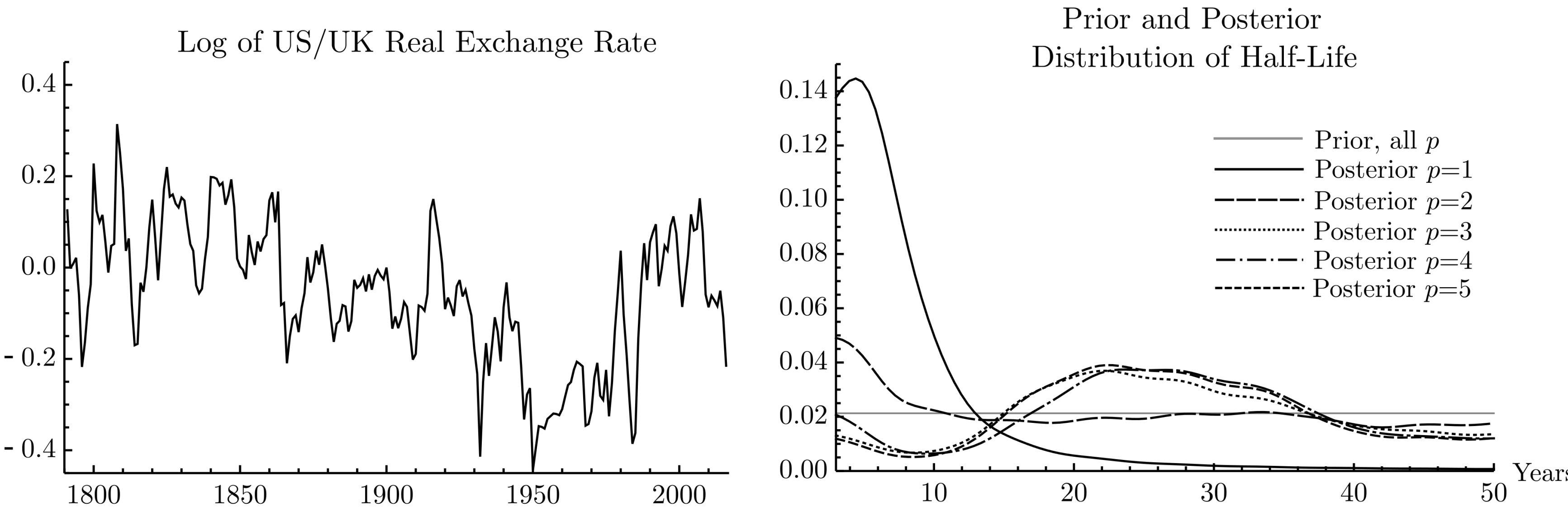
Persistence of Long-Run PPP Deviations

- Use GLTU(p) model for PPP deviations of US/UK real exchange rate

- For given value of p, Bayesian limited information inference for $N = 50$

- Prior also asymptotically important. Normalized such that for each p, prior on half-life in years is flat on $[3, 50]$

Posterior for US/UK Exchange Rate



$p \geq 1$ leads to substantially longer half-lives

Conclusions

- Flexible model for large sample persistence in economic time series

- Nearly unconstrained starting point for stationary processes that are not entirely different from benchmark I(1) model

- Frequentist inference difficult due to additional nuisance parameters that cannot be consistently estimated
  - But not a good reason to insist that such richer dynamics cannot exit
  - Potential constructive approach is Bayesian analysis with careful prior