Testing Models of Low-Frequency Variability

Ulrich K. Müller   Mark M. Watson
Princeton University

May 4, 2007
Motivation

• Pronounced interest in understanding low frequency behavior of economic time series

• Typical questions about low-frequency behavior

1. Is there a unit root?

2. What is the size of the largest AR root?

3. What is the value of \(d\) in a fractional model?

4. Is there cointegration, i.e. is \(x_t - y_t\) an \(I(0)\) series?
What is 'Low Frequency'? 

- Consider $I(0)$ model for a macroeconomic aggregate
  - $\Rightarrow$ roughly stationary, no strong persistence

- Expect lots of particular dynamics at business cycle frequencies (and higher)

- Plausible definition of $I(0)$ model for macro data: no additional fancy dynamics below business cycle frequencies (spectrum approximately flat below business cycle frequency)

- Focus on below business cycle frequencies $=$ period greater than 8 years

- Similar arguments for other models
How Much Information about Low Frequencies?

Spectrum of quarterly data

- 50 years of data: about 7 periodogram ordinates

- Asymptotic approximations with \#ordinates\to\infty potentially misleading
Our Approach

1. Extract low-frequency information by computing \( q \) weighted averages of time series data, where weights are low frequency trigonometric series

2. Asymptotic analysis of properties of finite number of weighted averages in common models

3. Measure fit of model by comparing transformed data from (1) with implication of (2)

Plan of Talk

1. Introduction

2. Methodology
   (a) Common time series models
   (b) Asymptotic properties of weighted averages
   (c) Choice of weights
   (d) Tests

3. Empirical Results

4. Conclusion
Common Time Series Models

Data Generating Process for observed data \( \{y_t\}_{t=1}^{T} \)

\[ y_t = d_t + u_t \]

where \( d_t = \mu \) or \( d_t = \mu + \beta t \).

1. Local-to-Unity AR model with parameter \( c \) (OU):

\[ u_t = (1 - c/T)u_{t-1} + \eta_t \]

such that \( T^{-1/2} u_{[1,T]} \Rightarrow \omega J_c(\cdot) \), where \( dJ_c(s) = -cJ_c(s)ds + dW(s) \)

2. Local-Level Model with parameter \( g \geq 0 \) (LLM):

\[ u_t = w_t + \frac{g}{T} \sum_{s=1}^{t} \eta_s \]

such that \( T^{-1/2} \sum_{t=1}^{[T]} u_t \Rightarrow \omega W_1(\cdot) + \omega g \int_0^T W_2(l)dl \)
3. Stationary Fractional Model with parameter $-1/2 < d < 1/2$:

$$(1 - L)^d u_t = \eta_t$$

such that $T^{-1/2 - d} \sum_{t=1}^{[T]} u_t \Rightarrow \omega W^d(\cdot)$, where $W^d$ is a 'type I' fractional Wiener process

+ Integrated version of these models, so that $u_t - u_{t-1}$ is modelled as above
Asymptotic Properties of Weighted Averages

- All models satisfy

\[ T^{-\alpha} \sum_{t=1}^{\cdot T} u_t \Rightarrow \omega G(\cdot) \]  

(1)

for some \( \alpha \) and \( \omega \), where \( G \) is a mean-zero Gaussian process with a covariance kernel \( k(r, s) = E[G(r)G(s)] \) that depends on the model and its parameter.

- Deal with deterministic component \( d_t \) by basing analysis on OLS residuals \( \{u_t^i\} \) of a regression of \( \{y_t\} \) on \( \{1\} \) (\( i = \mu \)) and \( \{1, t\} \) (\( i = \tau \)), respectively. By standard OLS algebra and (1)

\[ T^{-\alpha} \sum_{t=1}^{\cdot T} u_t^i \Rightarrow \omega G^i(\cdot) \]

where the covariance kernel of \( G^i \) can be computed from \( k(r, s) \).
Asymptotic Properties of Weighted Averages

- Let \( \Psi(\cdot) = (\Psi_1(\cdot), \ldots, \Psi_q(\cdot))' \), where \( \Psi_l : [0, 1] \rightarrow \mathbb{R}, \ l = 1, \ldots, q \), are functions with continuous derivative \( \psi_l \).

- Let \( S^i_t = \sum_{s=1}^{t} u^i_s \) (so that \( S^i_T = 0 \)). If \( T^{-\alpha}S^i_{[\cdot,T]} \Rightarrow \omega G^i(\cdot) \) then

\[
X_T \equiv T^{-\alpha} \sum_{t=1}^{T} \Psi(t/T) u^i_t \\
= T^{-\alpha} S^i_T \Psi(1) - T^{-\alpha} \sum_{t=1}^{T} S^i_{t-1} (\Psi(t/T) - \Psi((t - 1)/T)) \\
\Rightarrow -\omega \int_{0}^{1} G^i(\lambda) \psi(\lambda) d\lambda \sim \mathcal{N}(0, \omega^2 \Sigma)
\]

where the \( q \times q \) matrix \( \Sigma \) depends on \( \Psi, k(r, s) \) and \( i = \mu, \tau \).
Self-Normalized Weighted Averages

- We found

\[ X_T \equiv T^{-\alpha} \sum_{t=1}^{T} \Psi(t/T)u_t \Rightarrow \mathcal{N}(0, \omega^2 \Sigma) \equiv X \]

with \( \Sigma \) known for a given model and parameter. But what about \( \omega \) (and \( \alpha \))? 

- Restrict attention to scale invariant inference based on \( X_T \). Maximal invariant is given by

\[ v_T = \frac{X_T}{\sqrt{X_T'X_T}} \Rightarrow \frac{X}{\sqrt{X'X}} = v \]

and the density of \( v \) only depends on \( \Sigma \) (in fact, only on \( \Sigma(d)/\text{tr} \Sigma(d) \))

\[ f_v(\Sigma) \propto |\Sigma|^{-1/2}(v'\Sigma^{-1}v)^{-q/2} \]
Note on Continuity for Fractional Model

• Definition of stationary fractional model for $-1/2 < d < 1/2$:

\[ T^{-1/2-d} \sum_{t=1}^{[T]} u_t \Rightarrow \omega W^d(\cdot) \]

For $1/2 < d < 3/2$: $u_t - u_{t-1}$ is stationary fractional model with parameter $d - 1$ (as in Velaso (1999)), so that

\[ T^{-1/2-d} \sum_{t=1}^{[T]} u_t \Rightarrow \omega \int_0^T W^{d-1}(l)dl \]

• Consider $\Sigma = \Sigma(d)$ in fractional model. It turns out that

\[ \frac{\Sigma(d)}{\text{tr } \Sigma(d)} \]

can be continuously extended at $d = 1/2$ for $i = \mu, \tau$, so that likelihood of $v$ becomes continuous function of $-1/2 < d < 3/2$. 

Choice of Weights

• Want to extract low-frequency information only.

• Consider $R^2$ from regression of generic periodic series

$$\sin(\pi rs + a)$$
on candidate $\Psi(\cdot) = (\Psi_1(\cdot), \cdots, \Psi_q(\cdot))'$.

• Ideally, $R^2 = 1$ for $r \leq r_0$ and $R^2 = 0$ for $r > r_0$ for all $a$, where $r_0$ is the business-cycle cut-off frequency.

• Choice of $\Psi_l(s) = \sqrt{2}\cos(\pi ls)$, $l = 1, \cdots, q$, comes reasonably close to this ideal with $r_0 = q$. 
$R^2$ as Function of $r$ for $q = 14$

demeaned case $i = \mu$

average over $a$  
maximum over $a$  
minimum over $a$
• With $\Psi_l(s) = \sqrt{2} \cos(\pi ls)$, $\Sigma = I_q$ in $I(0)$ model and exactly diagonal in unit root model.

• Obtain same result for detrended case $i = \tau$ by choosing $\Psi_l$ as the eigenfunctions of the covariance kernel of detrended Wiener process.

• This choices lead to $\Sigma$ that are close to diagonal for all models and relevant parameter values (average absolute correlation with $q = 14$ is $<0.03$ for all models)
Square Root of Diagonal Elements of $\Sigma$

Local Level

Local-to-Unity
Tests

- For each model and parameter value, $X_T \Rightarrow \mathcal{N}(0, \omega^2 \Sigma)$, which implies specific asymptotic distribution for

$$v_T = \frac{X_T}{\sqrt{X_T^T X_T}}$$

that only depends on $\Sigma$.

- Observe $v_T$. Is it compatible with a specific model and parameter?
Test of Heteroskedasticity in $X_T$

- Let $X \sim \mathcal{N}(0, \Sigma)$ and $v = X/\sqrt{X'X}$. Consider test of

  $H_0 : \Sigma = \Sigma_0$ against $H_1 : \Sigma = \Lambda \Sigma_0 \Lambda$

where $\Lambda = \text{diag}(\exp(\delta_1), \cdots, \exp(\delta_q))$, and $\delta \sim \mathcal{N}(0, \gamma^2 \Omega)$.

  $\Rightarrow$ Tests the implication for the variance of $X_T$ of the various models of persistence for $u_t$ against a more flexible alternative

- We derive locally best test statistic (as $\gamma \to 0$) for $\delta = (\delta_1, \cdots, \delta_q)'$ a demeaned random walk, denoted LBIM.
Test of Low-Frequency Heteroskedasticity in $u_t$

- Models imply time invariant long-run variance. For instance, I(0) model for $u_t$ is defined as

$$T^{-1/2} \sum_{t=1}^{[T]} u_t \Rightarrow \omega \int_0^l dW(l)$$  \hspace{1cm} (2)

- Alternative I(0) model with time varying long-run variance $h(\cdot)$:

$$T^{-1/2} \sum_{t=1}^{[T]} u_t \Rightarrow \int_0^l h(l) dW(l)$$

Under this model, $X_T = T^{-1/2} \sum_{t=1}^T \Psi(t/T) u^i_t \Rightarrow \mathcal{N}(0, \Sigma(h))$ with $\Sigma$ a function of $h$ (and $h(\cdot) = \omega$ recovers (2))
Test of Low-Frequency Heteroskedasticity in $u_t$

• We consider weighted average power maximizing tests $H$, where the weight for alternative long-run variance paths $h(\cdot)$ is the distribution of $\exp[\kappa W^*(\cdot)]$

• For more general models than the I(0) model, the alternative model with low-frequency heteroskedasticity has time varying variances of the in-sample part of the "natural" MA representation of the asymptotic model

• Example: Stationary Ornstein-Uhlenbeck model

$$J_c(s) = \int_{-\infty}^{0} e^{-c(s-l)}dW(l) + \int_{0}^{s} e^{-c(s-l)}h(l)dW(l)$$
Likelihood Ratio Tests

- With $X \sim \mathcal{N}(0, \Sigma)$, best scale invariant test statistic to distinguish $H_0 : \Sigma = \Sigma_0$ against $H_1 : \Sigma = \Sigma_1$ is

$$LR = \frac{v'\Sigma_1^{-1}v}{v'\Sigma_0^{-1}v} = \frac{X'\Sigma_1^{-1}X}{X'\Sigma_0^{-1}X}$$

- Applications:

1. Point-optimal test LFUR of unit root model against local-to-unity alternatives with $c = 7.5$ in mean case and $c = 13.5$ in trend case, analogous to Elliott, Rothenberg, Stock (1996)

2. Point-optimal test LFST of $I(0)$ model against Local Level Model with $g = 8$ and $g = 13$ in mean and trend case, similar to Nyblom (1989) and KPSS (1992)
Discrimination Between Models?

- Quantify the difficulty using Total Variation Distance

  - The total variation distance between two probability measures $P_f$ and $P_g$ is

    $$\text{TVD}(f, g) = \sup_A |P_f(A) - P_g(A)|$$

    over all Borel sets $A$.

  - If $f$ and $g$ are two densities with respect to a common dominating measure $\mu$, then

    $$\text{TVD}(f, g) = \int 1[f < g](g - f)\,d\mu = \int 1[\frac{f}{g} < 1](1 - \frac{f}{g})g\,d\mu$$
TVD Between Models for $q = 14$, mean case

- Fractional
  - OU
  - LLM

- Ornstein-Uhlenbeck
  - FR
  - LLM

- Local Level Model
  - FR
  - OU
Empirical Analysis

We consider 21 US macroeconomic and financial data series

- Postwar quarterly macroeconomic series (GDP, interest rates, inflation, income/consumption)

- Annual long series (GNP, inflation, real exchange rates, price/earnings ratio)

- Daily S&P500 absolute returns 1928-2005
Four questions

1. Is the unit root model ($d = 1$ in fractional model, $c = 0$ in OU model, $g = 0$ in integrated LLM) consistent with data?

2. Is the $I(0)$ model ($d = 0$ in fractional model, $g = 0$ in LLM) consistent with the data?

3. Are there entire classes of models that are rejected?

4. How do the results of this analysis compare to results obtained from standard methods?
Postwar Real GDP \((q = 13)\)
Postwar Real GDP \((q = 13)\)

\[ \begin{array}{c|ccccc}
\text{p-value} & \text{LFUR} & \text{LFST} & \text{DF-GLS} & \text{KPSS} \\
\hline
0.34 & 0.01 & 0.16 & 0.00 \\
\end{array} \]

\( \Rightarrow \) \(I(0)\) model rejected, unit root not rejected
Postwar Labor Productivity ($q = 13$)
Postwar Labor Productivity \((q = 13)\)

\[
\begin{array}{c|ccccc}
& \text{LFUR} & \text{LFST} & \text{DF-GLS} & \text{KPSS} \\
p-value & 0.94 & 0.00 & 0.84 & 0.00 \\
\end{array}
\]

⇒ More persistence than I(1) model
Real Bond Rates 1900-2004 ($q = 26$)
Real Bond Rates 1900-2004 \( (q = 26) \)

\[
\begin{array}{c|cccc}
\text{LFUR} & \text{LFST} & \text{DF-GLS} & \text{KPSS} \\
p-value & 0.00 & 0.16 & 0.00 & 0.21 \\
\end{array}
\]

⇒ All models rejected due to second moment instability
Postwar Consumption/Income Ratio ($q = 13$)
Postwar Consumption/Income Ratio \(( q = 13 \) )

\[ (q = 13) \]

<table>
<thead>
<tr>
<th>p-value</th>
<th>LFUR</th>
<th>LFST</th>
<th>DF-GLS</th>
<th>KPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.85</td>
<td>0.00</td>
<td>0.91</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

⇒ \( I(0) \) model (=cointegration) rejected
Postwar 10 Year–1 Year Interest Spread ($q = 13$)
Postwar 10 Year–1 Year Interest Spread ($q = 13$)

<table>
<thead>
<tr>
<th>p-value</th>
<th>LFUR</th>
<th>LFST</th>
<th>DF-GLS</th>
<th>KPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.19</td>
<td>0.00</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

⇒ $I(0)$ model (=cointegration) not rejected, but second moment instability
Real Exchange Rates US-UK 1780-1992 \((q = 48)\)
Real Exchange Rates US-UK 1780-1992 ($q = 48$)

$\Rightarrow I(0)$ model rejected, unit root rejected, fractional model and local level model fit well
Daily Absolute Returns 1927-2004 ($q = 18$)
Daily Absolute Returns 1927-2004 \((q = 18)\)

<table>
<thead>
<tr>
<th></th>
<th>LFUR</th>
<th>LFST</th>
<th>DF-GLS</th>
<th>KPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

\(\Rightarrow I(0)\) model rejected, unit root model rejected, fractional model and local level model fit best, second moment instability
Conclusions

• Theoretical results
  1. Method to assess model fit of standard models at low frequencies
  2. Quantification of difficulty of distinguishing low-frequency models

• Empirical results
  1. $I(0)$ model is mostly rejected, even for putative cointegration relationships
  2. Unit root model fares much better
  3. Fractional model fits better for some series that are typically modelled as autoregressions
  4. For some series, too much heteroskedasticity in the underlying data for all models