Testing for Granger Causality with Mixed Frequency Data

Eric Ghysels[1], Jonathan B. Hill[1], & Kaiji Motegi[2]

[1]University of North Carolina at Chapel Hill
[2]Waseda University

September 15, 2014

Introduction

- Time series are often sampled at different frequencies,
  e.g. daily, monthly, quarterly, etc.
- Classic multivariate time series analysis is designed for single-frequency data.
  - $\Rightarrow$ Temporal aggregation of high frequency variables into the common lowest frequency.
  - $\Rightarrow$ Inaccurate statistical inference.
- **How can we exploit all data available whatever their sampling frequencies are?**

Introduction

- Mixed Data Sampling (MIDAS) econometrics.
  - Ghysels, Santa-Clara, and Valkanov (2004, WP).
  - Ghysels, Santa-Clara, and Valkanov (2006, JoE).
  - Andreou, Ghysels, and Kourtellos (2010, JoE).
- Ghysels' (2012, WP) mixed frequency vector autoregression (MF-VAR).

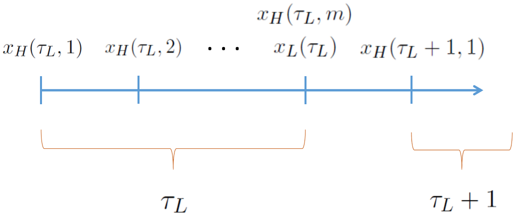
  – VAR model for mixed frequency data.

Introduction

- Based on Ghysels' MF-VAR, we develop Granger causality tests for mixed frequency data (henceforth "MF causality test").
- MF causality test achieves **higher local asymptotic power** than existing single-frequency tests do.
- In empirical application, MF causality test yields **more intuitive** results than existing single-frequency tests do.

Methodology

- x_L is a low frequency variable.
- x_H is a high frequency variable.
- In each low frequency time period τ_L, we sequentially observe $x_H(\tau_L, 1)$, $x_H(\tau_L, 2)$, $\dots$, $x_H(\tau_L, m)$, $x_L(\tau_L)$.



- Classic approach works on aggregated x_H but the present paper does not.

Methodology

- Instead of working on aggregated x_H, we stack all observations in each low frequency period τ_L:

$$\underbrace{\boldsymbol{X}(\tau_L)}_{\textit{mixed frequency vector}} = \begin{bmatrix} x_H(\tau_L, 1) \\ \vdots \\ x_H(\tau_L, m) \\ x_L(\tau_L) \end{bmatrix}.$$

- Ghysels' (2012) MF-VAR model assumes that $\boldsymbol{X}(\tau_L)$ itself follows VAR(q):

$$\boldsymbol{X}(\tau_L) = \sum_{k=1}^{q} \boldsymbol{A}_k \boldsymbol{X}(\tau_L - k) + \boldsymbol{\epsilon}(\tau_L).$$

Methodology

$$\underbrace{\begin{bmatrix} x_H(\tau_L,1) \\ \vdots \\ x_H(\tau_L,m) \\ x_L(\tau_L) \end{bmatrix}}_{=\boldsymbol{X}(\tau_L)} = \sum_{k=1}^{q} \underbrace{\begin{bmatrix} \boldsymbol{A}_{HH,k} & \boldsymbol{a}_{HL,k} \\ \boldsymbol{a}_{LH,k} & a_{LL,k} \end{bmatrix}}_{=\boldsymbol{A}_k} \underbrace{\begin{bmatrix} x_H(\tau_L-k,1) \\ \vdots \\ x_H(\tau_L-k,m) \\ x_L(\tau_L-k) \end{bmatrix}}_{=\boldsymbol{X}(\tau_L-k)} + \boldsymbol{\epsilon}(\tau_L).$$

- x_H does not Granger cause x_L given mixed frequency information set $\Leftrightarrow \boldsymbol{a}_{LH,1} = \cdots = \boldsymbol{a}_{LH,q} = \boldsymbol{0}_{1\times m}$.
- x_L does not Granger cause x_H given mixed frequency information set $\Leftrightarrow \boldsymbol{a}_{HL,1} = \cdots = \boldsymbol{a}_{HL,q} = \boldsymbol{0}_{m\times 1}$.
- These zero restrictions can be tested via usual asymptotics, e.g. Wald tests with χ^2_{mq}.

Local Asymptotic Power Analysis

- We show that the MF causality test achieves higher local asymptotic power than the classic low frequency causality test does.
- Suppose that the true data generating process (DGP) is a bivariate high frequency VAR (HF-VAR) of order 1:

$$\begin{bmatrix} x_H(\tau_L, j) \\ x_L(\tau_L, j) \end{bmatrix} = \begin{bmatrix} \phi_{HH} & \nu/\sqrt{T_L} \\ 0 & \phi_{LL} \end{bmatrix} \begin{bmatrix} x_H(\tau_L, j-1) \\ x_L(\tau_L, j-1) \end{bmatrix} + \begin{bmatrix} \eta_H(\tau_L, j) \\ \eta_L(\tau_L, j) \end{bmatrix}.$$

Local Asymptotic Power Analysis

$$\begin{bmatrix} x_H(\tau_L, j) \\ x_L(\tau_L, j) \end{bmatrix} = \begin{bmatrix} \phi_{HH} & \nu/\sqrt{T_L} \\ 0 & \phi_{LL} \end{bmatrix} \begin{bmatrix} x_H(\tau_L, j-1) \\ x_L(\tau_L, j-1) \end{bmatrix} + \begin{bmatrix} \eta_H(\tau_L, j) \\ \eta_L(\tau_L, j) \end{bmatrix}.$$

- x_H does not cause x_L given high frequency information set.
- x_L does cause x_H but with vanishing impact $\nu/\sqrt{T_L} \to 0$.
- $\nu \in \mathbb{R}$ is called the Pitman drift, representing the strength of causality from x_L to x_H.

Local Asymptotic Power Analysis

- Assume stock sampling $x_L(\tau_L) = x_L(\tau_L, m)$.
- The mixed frequency vector $\boldsymbol{X}(\tau_L)$ follows MF-VAR(1):

$$\underbrace{\begin{bmatrix} x_H(\tau_L,1) \\ \vdots \\ x_H(\tau_L,m) \\ x_L(\tau_L) \end{bmatrix}}_{=\boldsymbol{X}(\tau_L)} = \begin{bmatrix} 0 & \dots & 0 & \phi_{HH}^{1} & (\nu/\sqrt{T_L}) \sum_{j=1}^{1} \phi_{HH}^{1-j} \phi_{LL}^{j-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \phi_{HH}^{m} & (\nu/\sqrt{T_L}) \sum_{j=1}^{m} \phi_{HH}^{m-j} \phi_{LL}^{j-1} \\ 0 & \dots & 0 & 0 & \phi_{LL}^{m} \end{bmatrix} \underbrace{\begin{bmatrix} x_H(\tau_L-1,1) \\ \vdots \\ x_H(\tau_L-1,m) \\ x_L(\tau_L-1) \end{bmatrix}}_{=\boldsymbol{X}(\tau_L-1)} + \boldsymbol{\epsilon}(\tau_L).$$

- When we implement the mixed frequency causality test, the resulting Wald statistic follows $\chi_m^2(\kappa_{MF})$.
- Noncentrality parameter κ_{MF} can be characterized analytically.

Local Asymptotic Power Analysis

- Assume stock sampling:

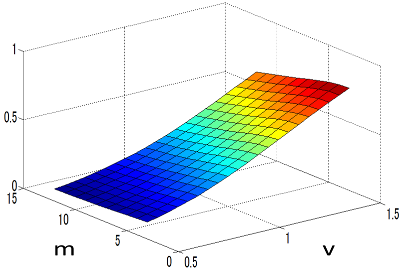
$$x_L(\tau_L) = x_L(\tau_L, m) \quad \text{and} \quad x_H(\tau_L) = x_H(\tau_L, m).$$

- The low frequency vector $\underline{\boldsymbol{X}}(\tau_L) = [x_H(\tau_L), x_L(\tau_L)]'$ follows VAR(1):

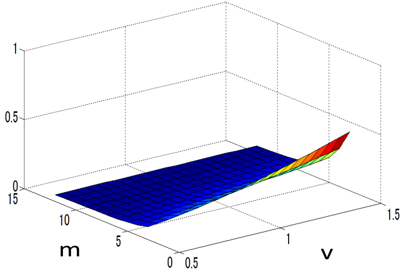
$$\underbrace{\begin{bmatrix} x_H(\tau_L) \\ x_L(\tau_L) \end{bmatrix}}_{=\underline{\boldsymbol{X}}(\tau_L)} = \begin{bmatrix} \phi_{HH}^m & (\nu/\sqrt{T_L}) \sum_{j=1}^m \phi_{HH}^{m-j} \phi_{LL}^{j-1} \\ 0 & \phi_{LL}^m \end{bmatrix} \underbrace{\begin{bmatrix} x_H(\tau_L - 1) \\ x_L(\tau_L - 1) \end{bmatrix}}_{=\underline{\boldsymbol{X}}(\tau_L - 1)} + \underline{\boldsymbol{\epsilon}}(\tau_L).$$

- Local asymptotic power can be computed analogously.

Local Asymptotic Power Analysis



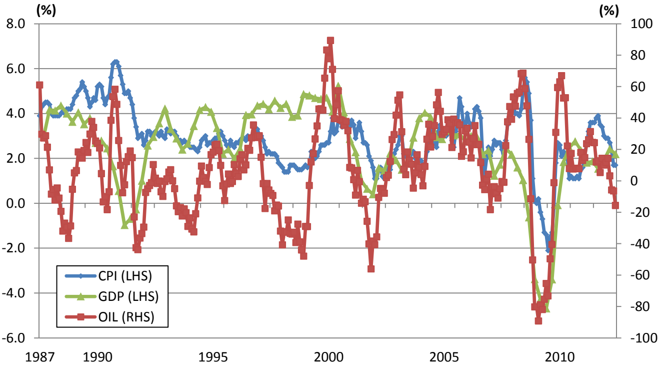
Mixed Frequency
(Original x_H & Aggregated x_L)



Low Frequency
(Aggregated x_H & Aggregated x_L)

Note: We assume $(\phi_{HH}, \phi_{LL}) = (0.25, 0.75)$, i.e. low persistence in x_H and high persistence in x_L. See the full paper for other parametrizations.

Empirical Application (U.S. Macroeconomy)



- **Monthly** consumer price index (CPI).
- **Monthly** oil prices (OIL).
- **Quarterly** gross domestic product (GDP).

Bootstrapped p-values of MF-VAR(1) and LF-VAR(4)

Mixed Frequency (**Monthly** CPI, **Monthly** OIL & Quarterly GDP)

Horizon	1	2	3	4	5
CPI$\nrightarrow$OIL	0.391	0.128	0.559	0.636	0.165
CPI$\nrightarrow$GDP	0.195	0.098°	**0.049**$^\circ$	0.100	0.180
OIL$\nrightarrow$GDP	0.680	0.548	0.236	0.300	0.196
OIL$\nrightarrow$CPI	**0.002**$^\circ$	0.182	0.439	**0.029**$^\circ$	0.605
GDP$\nrightarrow$CPI	**0.015**$^\circ$	0.570	0.583	0.125	0.500
GDP$\nrightarrow$OIL	0.724	0.833	0.895	0.855	0.946

Low Frequency (**Quarterly** CPI, **Quarterly** OIL & Quarterly GDP)

Horizon	1	2	3	4	5
CPI$\nrightarrow$OIL	**0.035**$^\circ$	0.095°	0.095°	0.116	0.492
CPI$\nrightarrow$GDP	0.380	0.215	0.272	0.238	0.683
OIL$\nrightarrow$GDP	0.145	**0.044**$^\circ$	0.088°	**0.027**$^\circ$	0.066°
OIL$\nrightarrow$CPI	0.206	0.320	0.986	0.710	0.521
GDP$\nrightarrow$CPI	0.680	0.497	0.323	0.596	0.645
GDP$\nrightarrow$OIL	0.095°	0.164	0.516	0.376	0.541

Conclusions

- **The mixed frequency Granger causality test** is a Wald test based on Ghysels' (2012) mixed frequency vector autoregression.
- The MF causality test has **higher local asymptotic power** than the LF causality test does.
- In empirical application the MF test and the LF test produce very different results, and **the MF test yields more intuitive causal implications**.

References

- Andreou, E., E. Ghysels, and A. Kourtellos (2010). Regression Models with Mixed Sampling Frequencies. *Journal of Econometrics*, 158, 246-261.
- Ghysels, E. (2012). Macroeconomics and the Reality of Mixed Frequency Data. Working Paper at the University of North Carolina at Chapel Hill.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. Working Paper at the University of North Carolina at Chapel Hill.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting Volatility: Getting the Most out of Return Data Sampled at Different Frequencies. *Journal of Econometrics*, 131, 59-95.