Online Appendix: Out-of-Sample Forecasting of Foreign Exchange Rates: The Band Spectral Regression and LASSO

Tatsuma Wada*

Keio University

This Version: March 7, 2020

Abstract

We propose to utilize the band spectral regression for out-of-sample forecasts of exchange rates. When one period ahead forecast is considered, there is some evidence that the band spectral regression benefits us especially when the Taylor rule fundamentals model is employed. However, when the forecasting horizon increases, the purchasing power parity (PPP) fundamentals model is found to be powerful and we can improve the out-of-sample forecast by selecting appropriate frequency bands. Bayesian model averaging shows that placing a high weight on the business cycle frequency improves the accuracy of the out-of-sample forecasting of the PPP model (as well as the monetary fundamentals model) when a longer forecasting horizon is our focus. Without specifying the frequency bands prior to applying the regression, LASSO can provide better out-of-sample exchange rate forecasts for many cases and provide information about the dynamic relationship between forecasting variables and exchange rates.

JEL Classification Number: C22; C52; F31; F47.

Keywords: Band Spectral Regression, Bayesian Model Averaging, Exchange Rate Models, Frequency Domain, LASSO.

*Faculty of Policy Management, Keio University, 5322 Endo, Fujisawa, Kanagawa, 252-0882, Japan (E-mail: twada@sfc.keio.ac.jp, Tel. +81-466-49-3451, Fax. +81-466-49-3451).

1 Online Appendix 1: The Band Spectral Regression and The Time-Domain OLS

Suppose process y is generated by

$$y_t = \sum_{j=-\infty}^{\infty} b_j^0 x_{t-j} + \varepsilon_t,$$

where $\{y_t, x_t\}$ are jointly stationary and $E[\varepsilon_t x_{t-j}] = 0$ for all j. Suppose that we run a regression assuming

$$y_t = \sum_{j=-\infty}^{\infty} b_j^1 x_{t-j} + u_t,$$

where $b_j^1 = 0$ for some j by assumption (hence, it is impossible to find $b_j^1 = b_j^0$ for all j.)

According to Sims (1972) and Sargent (1987), OLS is equivalent to finding b_j^1 by solving a minimization problem:

$$\min_{b_j^1} \int_{-\pi}^{\pi} \left| b^0\left(e^{-i\omega}\right) - b^1\left(e^{-i\omega}\right) \right|^2 g_x\left(e^{-i\omega}\right) d\omega, \tag{1}$$

where

$$b^k\left(e^{-i\omega}\right) = \sum_{j=-\infty}^{\infty} b_j^k e^{-ij\omega},$$

for $k = 0, 1$; and $g_x\left(e^{-i\omega}\right)$ is the spectral density of x_t.

Now, suppose that we estimate the model

$$y_t = b^1 x_t + u_t. \tag{2}$$

Then, the time domain regression involves minimizing

$$\int_{-\pi}^{\pi} \left| b^0\left(e^{-i\omega}\right) - b^1 \right|^2 g_x\left(e^{-i\omega}\right) d\omega.$$

The first-order-condition yields

$$2\widehat{b}_{OLS}^1 - \int_{-\pi}^{\pi} \left| b^0\left(e^{-i\omega}\right) + b^0\left(e^{i\omega}\right) \right| g_x\left(e^{-i\omega}\right) d\omega = 0,$$

giving the estimate of b^1, $\widehat{b}^1_{OLS}$ as

$$\widehat{b}^1_{OLS} = \frac{1}{2}\left[\int_{-\pi}^{\pi}\left(b^0\left(e^{-i\omega}\right) + b^0\left(e^{i\omega}\right)\right) g_x\left(e^{-i\omega}\right) d\omega\right].$$

Now, let us consider the band spectral regression. Suppose that we allow low frequency band $[0, \omega_0]$ and high frequency band $(\omega_0, \pi]$ to have different coefficients. More specifically, we assume β_A for the low frequency band and $\beta_{A^C}$for the high frequency band.

Then, we have

$$b^1\left(e^{-i\omega}\right) = \beta_A 1_{(|\omega|<\omega_0)} + \beta_{A^C} 1_{(|\omega|>\omega_0)},$$

which, in the case of our regression model (2) means that,

$$b^1 = \frac{\beta_A \omega_0}{\pi} + \frac{\beta_{A^C}}{\pi}\left(\pi - \omega_0\right).$$

Because band spectral regression minimizes the squared residuals in the time domain (see Corbae et al. 2002, among others) to find the unknown parameters β_A and β_{A^C}, the minimization problem is the same as (1) . Henceforth,

$$\begin{aligned}
& \min_{\beta_A, \beta_{A^C}} \int_{-\pi}^{\pi} \left|b^0\left(e^{-i\omega}\right) - \beta_A 1_{(|\omega|<\omega_0)} - \beta_{A^C} 1_{(|\omega|>\omega_0)}\right|^2 g_x\left(e^{-i\omega}\right) d\omega \\
= \; & \min\left\{\int_{\omega_0}^{\pi} \left|b^0\left(e^{-i\omega}\right) - \beta_{A^C}\right|^2 g_x\left(e^{-i\omega}\right) d\omega + \int_{-\omega_0}^{\omega_0} \left|b^0\left(e^{-i\omega}\right) - \beta_A\right|^2 g_x\left(e^{-i\omega}\right) d\omega \right. \\
& \left. + \int_{-\pi}^{-\omega_0} \left|b^0\left(e^{-i\omega}\right) - \beta_{A^C}\right|^2 g_x\left(e^{-i\omega}\right) d\omega \right\}
\end{aligned}$$

It is clear that the band spectral regression can generally yield a smaller mean squared error because it has two free parameters to be estimated. However, it is not immediately clear whether the band spectral regression with a restriction that $\beta_{A^C} = 0$ prevails in the time domain regression. To see this, set $\beta_{A^C} = 0$, and then consider the minimization problem:

$$\begin{aligned}
& \min_{\beta_A}\left\{\int_{\omega_0}^{\pi} \left|b^0\left(e^{-i\omega}\right)\right|^2 g_x\left(e^{-i\omega}\right) d\omega + \int_{-\omega_0}^{\omega_0} \left|b^0\left(e^{-i\omega}\right) - \beta_A\right|^2 g_x\left(e^{-i\omega}\right) d\omega \right. \\
& \left. + \int_{-\pi}^{-\omega_0} \left|b^0\left(e^{-i\omega}\right)\right|^2 g_x\left(e^{-i\omega}\right) d\omega \right\}.
\end{aligned}$$

The first order condition is

$$2\widehat{\beta}_A \int_{-\omega_0}^{\omega_0} g_x\left(e^{-i\omega}\right) d\omega - \int_{-\omega_0}^{\omega_0} \left(b^0\left(e^{-i\omega}\right) + b^0\left(e^{i\omega}\right)\right) g_x\left(e^{-i\omega}\right) d\omega = 0.$$

Hence,

$$\widehat{\beta}_A = \frac{\int_{-\omega_0}^{\omega_0} \left(b^0\left(e^{-i\omega}\right) + b^0\left(e^{i\omega}\right)\right) g_x\left(e^{-i\omega}\right) d\omega}{2\int_{-\omega_0}^{\omega_0} g_x\left(e^{-i\omega}\right) d\omega}.$$

Having two estimates, we can now compare the mean squared errors. The mean squared error for the band spectral regression is smaller than that for the time-domain OLS when:

$$\begin{aligned}
& \int_{\omega_0}^{\pi} \left|b^0\left(e^{-i\omega}\right)\right|^2 g_x\left(e^{-i\omega}\right) d\omega + \int_{-\omega_0}^{\omega_0} \left|b^0\left(e^{-i\omega}\right) - \widehat{\beta}_A\right|^2 g_x\left(e^{-i\omega}\right) d\omega + \int_{-\pi}^{-\omega_0} \left|b^0\left(e^{-i\omega}\right)\right|^2 g_x\left(e^{-i\omega}\right) d\omega \\
< \; & \int_{\omega_0}^{\pi} \left|b^0\left(e^{-i\omega}\right) - \widehat{b}^1_{OLS}\right|^2 g_x\left(e^{-i\omega}\right) d\omega + \int_{-\omega_0}^{\omega_0} \left|b^0\left(e^{-i\omega}\right) - \widehat{b}^1_{OLS}\right|^2 g_x\left(e^{-i\omega}\right) d\omega \\
& + \int_{-\pi}^{-\omega_0} \left|b^0\left(e^{-i\omega}\right) - \widehat{b}^1_{OLS}\right|^2 g_x\left(e^{-i\omega}\right) d\omega. \qquad (3)
\end{aligned}$$

To simplify the argument, let us assume that x_t is a white noise process, that is, $g_x\left(e^{-i\omega}\right) = 1/2\pi$. The estimates of the time domain OLS and the band spectral regression are then:

$$\begin{aligned}
\widehat{b}^1_{OLS} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} b^0\left(e^{-i\omega}\right) d\omega \text{ and} \\
\widehat{\beta}_A &= \frac{\int_{-\omega_0}^{\omega_0} b^0\left(e^{-i\omega}\right) d\omega}{2\omega_0},
\end{aligned}$$

respectively.

Then, condition (3) becomes:

$$
\begin{aligned}
&\frac{1}{2\pi}\int_{\omega_0}^{\pi}\left|b^0\left(e^{-i\omega}\right)\right|^2 d\omega+\frac{1}{2\pi}\int_{-\omega_0}^{\omega_0}\left|b^0\left(e^{-i\omega}\right)-\widehat{\beta}_A\right|^2 d\omega+\frac{1}{2\pi}\int_{-\pi}^{-\omega_0}\left|b^0\left(e^{-i\omega}\right)\right|^2 d\omega\\
<\;&\frac{1}{2\pi}\int_{\omega_0}^{\pi}\left|b^0\left(e^{-i\omega}\right)-\widehat{b}^1_{OLS}\right|^2 d\omega+\frac{1}{2\pi}\int_{-\omega_0}^{\omega_0}\left|b^0\left(e^{-i\omega}\right)-\widehat{b}^1_{OLS}\right|^2 d\omega\\
&+\frac{1}{2\pi}\int_{-\pi}^{-\omega_0}\left|b^0\left(e^{-i\omega}\right)-\widehat{b}^1_{OLS}\right|^2 d\omega
\end{aligned}
$$

The left hand side is

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|b^0\left(e^{-i\omega}\right)\right|^2 d\omega-\frac{1}{4\pi\omega_0}\left[\int_{-\omega_0}^{\omega_0} b^0\left(e^{-i\omega}\right)d\omega\right]^2,
$$

and the right hand side is

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|b^0\left(e^{-i\omega}\right)\right|^2 d\omega-\frac{1}{4\pi^2}\left[\int_{-\pi}^{\pi} b^0\left(e^{-i\omega}\right)d\omega\right]^2.
$$

Therefore, inequality (3) holds when

$$
-\frac{1}{4\pi\omega_0}\left[\int_{-\omega_0}^{\omega_0} b^0\left(e^{-i\omega}\right)d\omega\right]^2<-\frac{1}{4\pi^2}\left[\int_{-\pi}^{\pi} b^0\left(e^{-i\omega}\right)d\omega\right]^2
$$

or

$$
\frac{\left[\int_{-\omega_0}^{\omega_0} b^0\left(e^{-i\omega}\right)d\omega\right]^2}{\left[\int_{-\pi}^{\pi} b^0\left(e^{-i\omega}\right)d\omega\right]^2}>\frac{\omega_0}{\pi} \tag{4}
$$

The band spectral regression yields a smaller mean squared error as long as (4) is satisfied. Of course, the imprecise estimation of $\widehat{\beta}_A$ may lead to a larger estimated mean squared error (or average squared residual) for the band spectral regression than that for the time domain OLS.

2 Appendix 2: The LASSO Specification and Estimation

2.1 The LASSO Specification

The LASSO specification is

$$\begin{aligned} y_c &= \Psi X_c \beta_A + u_c \\ &= P\zeta + u_c. \end{aligned} \tag{5}$$

First, we define some vectors and matrices that are used to render $\Psi X_c \beta_A$ to the product of the known matrix P and the unknown vector ζ.

Let

$$a = vec\,(A)$$

and $\alpha = \begin{pmatrix} a_{11} & a_{22} & \cdots & a_{TT} \end{pmatrix}'$ are the diagonal elements of A. Then, it is well known that there exists a matrix C_d

$$\underbrace{C_d}_{T\times T^2} = diag \begin{pmatrix} e_1' & e_2' & \cdots & e_T' \end{pmatrix},$$

where $e_i = \begin{pmatrix} 0 & \cdots & \underbrace{1}_{i-th} & \cdots & 0 \end{pmatrix}'$, such that $C_d C_d' = I$ and

$$\alpha = C_d a \;\text{ and } a = C_d' \alpha.$$

Note further that

$$\begin{aligned} \alpha &= D_o \widetilde{\alpha}_o \;\text{ if T is odd} \\ &= D_e \widetilde{\alpha}_e \;\text{ if T is even,} \end{aligned}$$

where

$$
\underbrace{\widetilde{\alpha}_o}_{(T+1)/2\times 1} = \begin{bmatrix} a_{11} & a_{22} & a_{33} & \cdots & a_{(T+1)/2,(T+1)/2} \end{bmatrix}'
$$

$$
\underbrace{\widetilde{\alpha}_e}_{(T+2)/2\times 1} = \begin{bmatrix} a_{11} & a_{22} & a_{33} & \cdots & a_{(T+2)/2,(T+2)/2} \end{bmatrix}';
$$

$$
D_o = \underbrace{\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \\ & & 0 & 1 \\ & & 1 & \\ 0 & 1 & & \end{bmatrix}}_{T\times(T+1)/2}, \; D_e = \underbrace{\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & 0 \\ 0 & & & 1 \\ 0 & & \diagup & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{T\times(T+2)/2}
$$

$$
D_o = \begin{bmatrix} 1 & 0_{1\times(T-1)/2} \\ 0_{(T-1)/2\times 1} & I_{(T-1)/2} \\ 0_{(T-1)/2\times 1} & \widetilde{I}_{(T-1)/2} \end{bmatrix}, \; D_e = \underbrace{\begin{bmatrix} 1 & 0_{1\times(T-2)/2} & 0 \\ 0_{(T-2)/2\times 1} & I_{(T-2)/2} & 0_{(T-2)/2\times 1} \\ 0 & 0_{1\times(T-2)/2} & 1 \\ 0_{(T-2)/2\times 1} & \widetilde{I}_{(T-2)/2} & 0_{(T-2)/2\times 1} \end{bmatrix}}_{T\times(T+2)/2},
$$

and

$$\widetilde{I} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The reason why we define $\widetilde{\alpha}_o$ and $\widetilde{\alpha}_e$ is that the frequency bands $(\pi - h)$ and $(\pi + h)$ should have the same information. Hence, if $a_{(T+1)/2-k,(T+1)/2-k} = 1$ then $a_{(T+3)/2+k,(T+3)/2+k}$ must be 1 when T is odd. Similarly, $a_{(T+2)/2-k,(T+2)/2-k}$ and $a_{(T+2)/2+k,(T+2)/2+k}$ must have the same number (either 0 or 1) when T is even.

By using the vectors and matrices defined above, we render the first term on the right hand side:

$$\begin{aligned} \Psi X_c \beta_A &= vec\,(\Psi X_c \beta_A) \\ &= \left(\beta_A' \otimes \Psi\right) vec\,(X_c) \\ &= \left(vec\,(X_c)' \otimes I_T\right) vec\left(\beta_A' \otimes W'AW\right) \\ &= \left(vec\,(X_c)' \otimes I_T\right) \left(\beta_A \otimes \left(\overline{W}^T \otimes W'\right) C_d' \alpha\right) \\ &= \left(vec\,(X_c)' \otimes I_T\right) \left(\beta_A \otimes \left(\overline{W}^T \otimes W'\right) C_d' D_o \widetilde{\alpha}_o\right) \quad \text{if T is odd} \\ &= \left(vec\,(X_c)' \otimes I_T\right) \left(\beta_A \otimes \left(\overline{W}^T \otimes W'\right) C_d' D_e \widetilde{\alpha}_e\right) \quad \text{if T is even,} \end{aligned}$$

where $\overline{W}^T$ is a non-conjugate, transposed matrix of W.

Now, let us consider a case where β_A is scalar and a case where β_A is a $k \times 1$ vector. For simplicity, we only show the derivation of (5) when T is odd.

- When β_A is a scalar ($k = 1$)

$$\begin{aligned} &= \underbrace{\left(vec\,(X_c)' \otimes I_T\right)}_{T \times T^2} \underbrace{\left(\overline{W}^T \otimes W'\right)}_{T^2 \times T^2} \underbrace{C_d'}_{T^2 \times T} \underbrace{D_o}_{T \times (T+1)/2} \underbrace{\beta_A \widetilde{\alpha}_o}_{(T+1)/2 \times 1} \\ &= \underbrace{P}_{T \times T^2} \underbrace{\zeta}_{T^2 \times 1} \end{aligned}$$

where $P = \left(vec\left(X_c\right)' \otimes I_T\right)$,and $\zeta = \left(\overline{W}^T \otimes W'\right) C_d' D_o \beta_A \widetilde{\alpha}_o$. Therefore,

$$\zeta = \begin{bmatrix} \beta_{A,1} a_{11} & \cdots & \beta_{A,1} a_{(T+1)/2,(T+1)/2} \end{bmatrix}'.$$

- The estimation of A

$$\widehat{\beta_A \widetilde{\alpha}_o} = D_{h,o}' C_d \left(\overline{W} \otimes W\right) \widehat{\zeta}$$

where

$$D_{h,o} = \underbrace{\begin{bmatrix} 1 & & & 0 \\ & 1/2 & & \\ & & \ddots & \\ 0 & & & 1/2 \\ & & 0 & 1/2 \\ & & 1/2 & \\ 0 & 1/2 & & \end{bmatrix}}_{T \times (T+1)/2},$$

so that $D_{h,o}' D_o = I$.

- The Alternative LASSO specification

$$\begin{aligned}
\Psi X_c \beta_A &= \underbrace{\left(vec\left(X_c\right)' \otimes I_T\right)}_{T \times T^2} \underbrace{\left(\overline{W}^T \otimes W'\right)}_{T^2 \times T^2} \underbrace{C_d'}_{T^2 \times T} \underbrace{D_o}_{T \times (T+1)/2} \underbrace{\beta_A \widetilde{\alpha}_o}_{(T+1)/2 \times 1} \\
&= \underbrace{\left(vec\left(X_c\right)' \overline{W}^T \otimes W'\right)}_{T \times T^2} \underbrace{C_d'}_{T^2 \times T} \underbrace{D_o}_{T \times (T+1)/2} \underbrace{\beta_A \widetilde{\alpha}_o}_{(T+1)/2 \times 1} \\
&= \underbrace{P}_{T \times (T+1)/2} \underbrace{\zeta}_{(T+1)/2 \times 1},
\end{aligned}$$

where $P = \left(vec\left(X_c\right)' \overline{W}^T \otimes W'\right) C_d' D_o$, and $\zeta = \beta_A \widetilde{\alpha}_o$.

2.2 The Estimation of the Diagonal Elements of the A Matrix

Similar to the previous subsection, we consider two cases:

- If β_A is a scalar ($k = 1$)

  Since $\zeta = \beta_A \alpha_o$, we look at whether each row of ζ is non-zero by checking the absolute value of the i-th row to see that is less than 1e-5; that is $|\zeta_i| < 1e-5$. If so, we place 0 in the corresponding frequency, and place 1 otherwise.

- If β_A is a $k \times 1$ vector

  There are k estimates of the j-th frequency (multiplied by β_A): $\beta_A a_{j,j}$. If the minimum of the absolute value of k estimates is less than $1e-5$, we place 0 in the j-th frequency, and place 1 otherwise.

3 Appendix 3: The Inferences: Clark and McCracken (2012) Test

3.1 The Nested Model

$$\begin{aligned} y &= \Psi Z \left(\pi_1 - \Pi_2 \beta_A\right) + \Psi^c Z \left(\pi_1 - \Pi_2 \beta_{A^c}\right) + \Psi X \beta_A + \Psi^c X \beta_{A^c} + \varepsilon \\ &= \pi_1 - \Psi \Pi_2 \beta_A - \Psi^c \Pi_2 \beta_{A^c} + \Psi X \beta_A + \Psi^c X \beta_{A^c} + \varepsilon \end{aligned}$$

because $Z = 1$.

- If $\beta_{A^c} = 0$ (restricted), then the model is

  $$y = \pi_1 - \Psi \Pi_2 \beta_A + \Psi X \beta_A + \varepsilon.$$

  The null hypothesis is $\beta_A = 0$.

- If $\beta_{A^c} \neq 0$ (nonrestricted), then the model is

  $$y = \pi_1 - \Psi \Pi_2 \beta_A - \Psi^c \Pi_2 \beta_{A^c} + \Psi X \beta_A + \Psi^c X \beta_{A^c} + \varepsilon$$

The null hypothesis is $\beta_A = 0$ and $\beta_{A^c} = 0$.

In both cases, the model under the alternative hypothesis nests the model under the null hypothesis.

3.2 The Bootstrap Procedure

Clark and McCracken (2012) and summarized in Clark and McCracken (2013)

1. Step 1: Run a regression with the full set of regressors (unrestricted; $x'_{1,s}$) to obtain the residuals $\widehat{v}_{1,s+\tau}$
$$y_{t+s} = x'_{1,s}\widehat{\beta}_{1,T} + \widehat{v}_{1,s+\tau} \text{ for } s = 1, 2, \ldots, T - \tau.$$
Here, we treat $x'_{1,s}$ as the regressor using all frequencies, $(0, \pi)$ (i.e., we assume that y_{t+s} is generated from the time-domain regression model).

2. Step 2 Run a regression with the subset of regressors (restricted; $x'_{2,s}$) to obtain the residuals $\widehat{v}_{2,s+\tau}$
$$y_{t+s} = x'_{2,s}\widehat{\beta}_{2,T} + \widehat{v}_{2,s+\tau} \text{ for } s = 1, 2, \ldots, T - \tau.$$
In our case, $x'_{2,s} = 1$ because under the null hypothesis y_{t+s} follows the random walk process with a drift.

3. Step 3: For the unrestricted residuals, fit an MA$(\tau - 1)$ model provided $\tau > 1$.
$$\widehat{v}_{1,s+\tau} = \theta_1 \varepsilon_{s+\tau-1} + \theta_2 \varepsilon_{s+\tau-2} + \cdots + \theta_{\tau-1}\varepsilon_{s+1} + \varepsilon_{s+\tau}$$

4. Step 4: $\eta_s \sim i.i.d.N\,(0, 1)$ for $s = 1, 2, \ldots, T$ and
$$\begin{aligned} \widehat{v}^*_{1,s+1} &= \eta_{s+1}\widehat{v}_{1,s+1} \quad \text{for } \tau = 1 \\ \widehat{v}^*_{1,s+\tau} &= \widehat{\theta}_1\widehat{\varepsilon}_{s+\tau-1}\eta_{s+\tau-1} + \widehat{\theta}_2\widehat{\varepsilon}_{s+\tau-2}\eta_{s+\tau-2} + \cdots + \widehat{\theta}_{\tau-1}\widehat{\varepsilon}_{s+1}\eta_{s+1} + \widehat{\varepsilon}_{s+\tau}\eta_{s+\tau} \quad \text{for } \tau > 1 \end{aligned}$$

5. Step 5: Generate
$$y^*_{s+\tau} = x'_{2,s}\widehat{\beta}_{2,T} + \widehat{v}^*_{1,s+\tau}$$

6. Step 6: Using the artificial data $(y^*_{s+\tau})$, compute the test statistics

7. Step 7: Repeat Steps 4 to 6 N times. This gives the empirical distribution of the simulated statistics. Count the number of times the test statistic from the real data $(y_{s+\tau})$ exceed that from the artificial data $(y^*_{s+\tau})$ and divide it by N to get the p-value.

Note that under the null hypothesis, (the difference of) the exchange rate follows the random walk process, but under the alternative, the same y_{t+s} is used for all band spectral regressions (i.e., the regressions using high-, low-, middle-, business cycle- and all frequencies). Therefore, in each bootstrapping trial, the test statistics are computed from the same pseudo data for all the regressors employing different frequency bands.

3.3 Test Statistics

- MSE-F test:

$$F = \frac{\sum_{t=R+h+1}^{T} d_t}{(T-R-h)^{-1} \sum_{t=R+h+1}^{T} e_t^2} = (T-R-h)\left(\frac{1}{MSE\ ratio} - 1\right)$$

  where $d_t = re_t^2 - e_t^2$.

4 Figures

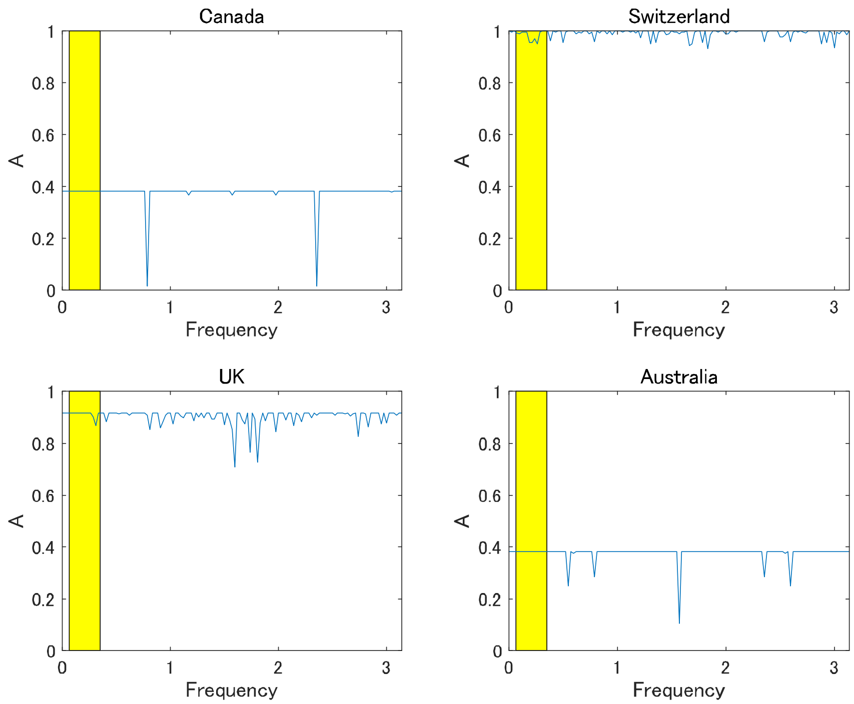

Figure 7 (a): Estimated Average A for the PPP fundamentals with $h = 1$.

(Business cycle frequencies are shaded.)

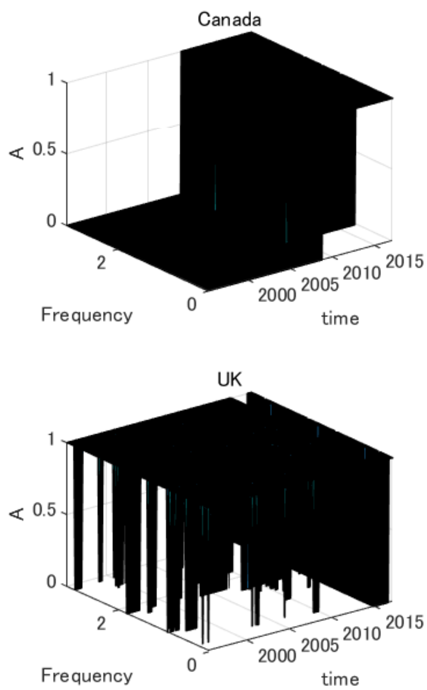

Figure 7 (b): Estimated A for the PPP fundamentals with $h = 1$.

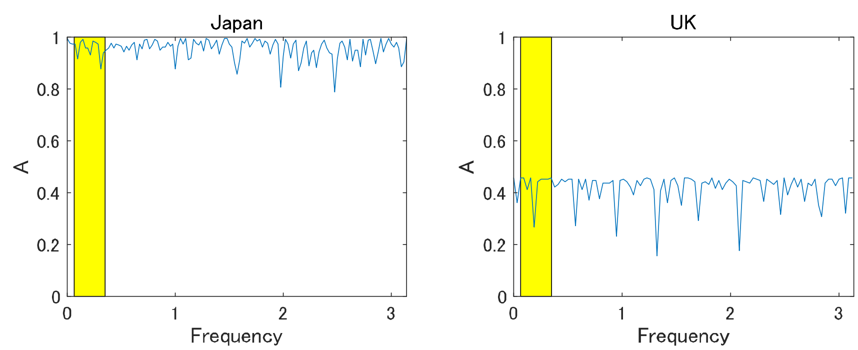

Figure 8(a): Estimated Average A for Monetary fundamentals with $h = 1$.

(Business cycle frequencies are shaded.)

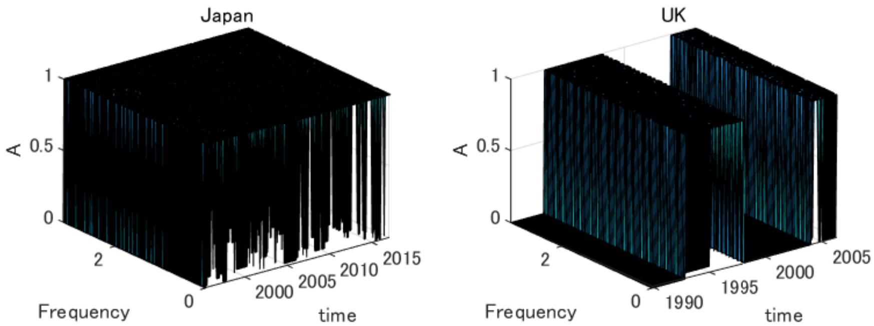

Figure 8(b): Estimated A for Monetary fundamentals with $h = 1$.

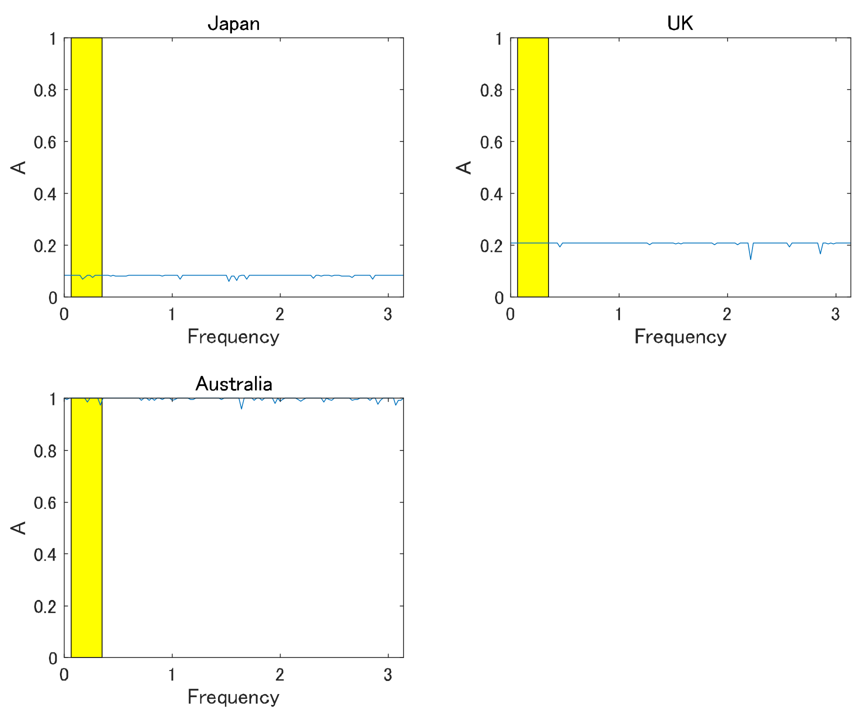

Figure 9(a): Estimated Average A for the Taylor Rule Fundamentals with $h = 1$.

(Business cycle frequencies are shaded.)

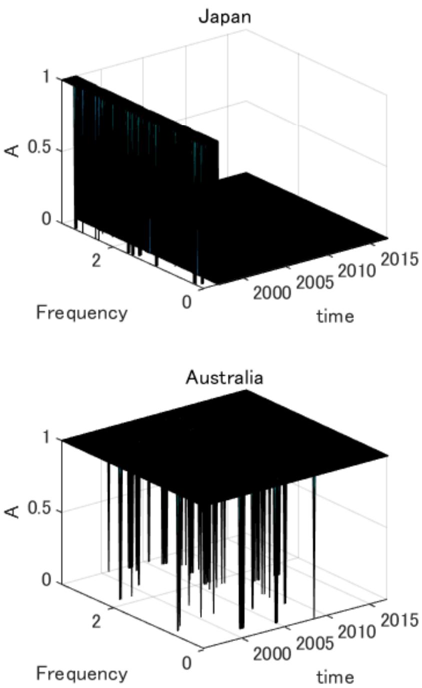

Figure 9(b): Estimated A for the Taylor Rule Fundamentals with $h = 1$.

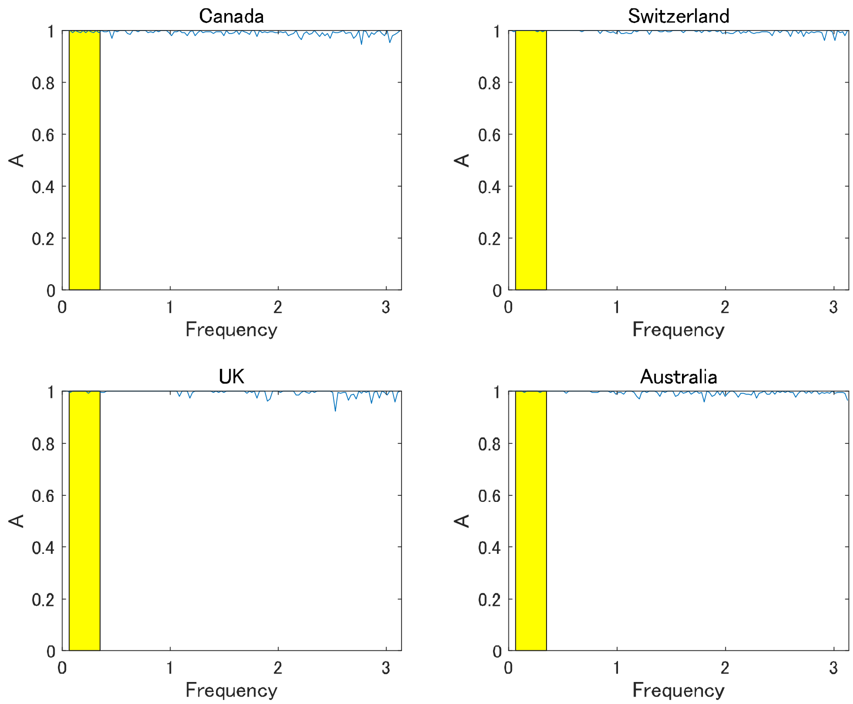

Figure 10(a): Estimated Average A for the PPP Fundamentals with $h = 6$.

(Business cycle frequencies are shaded.)

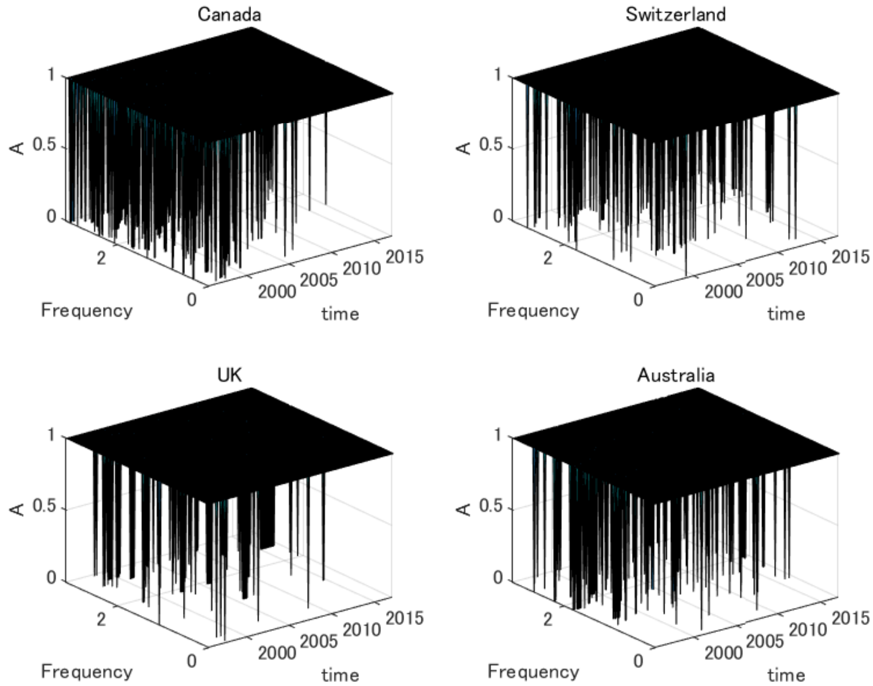

Figure 10(b): Estimated A for the PPP Fundamentals with $h = 6$.

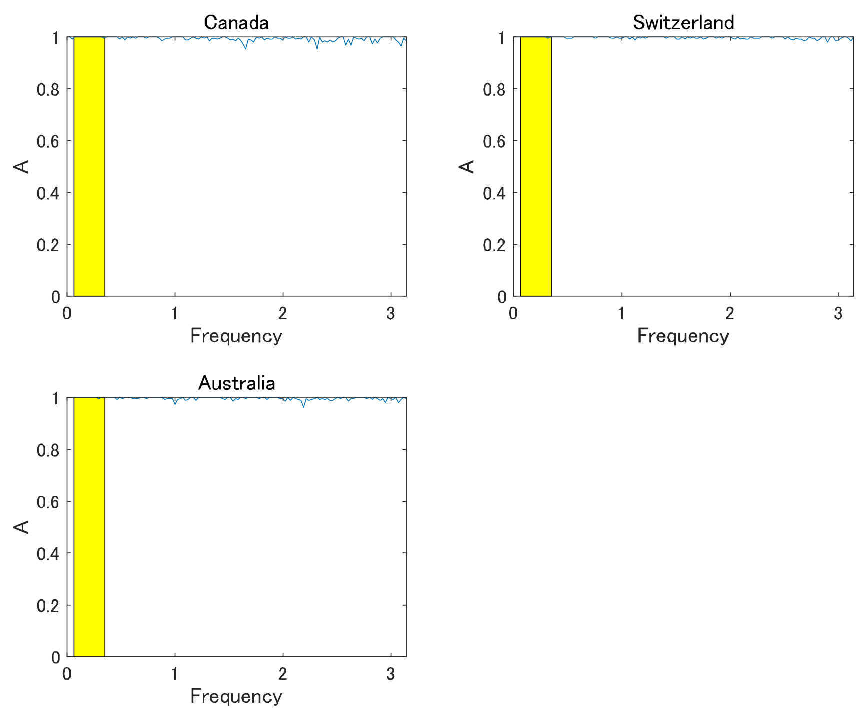

Figure 11(a): Estimated Average A for the PPP Fundamentals with $h = 12$.

(Business cycle frequencies are shaded.)

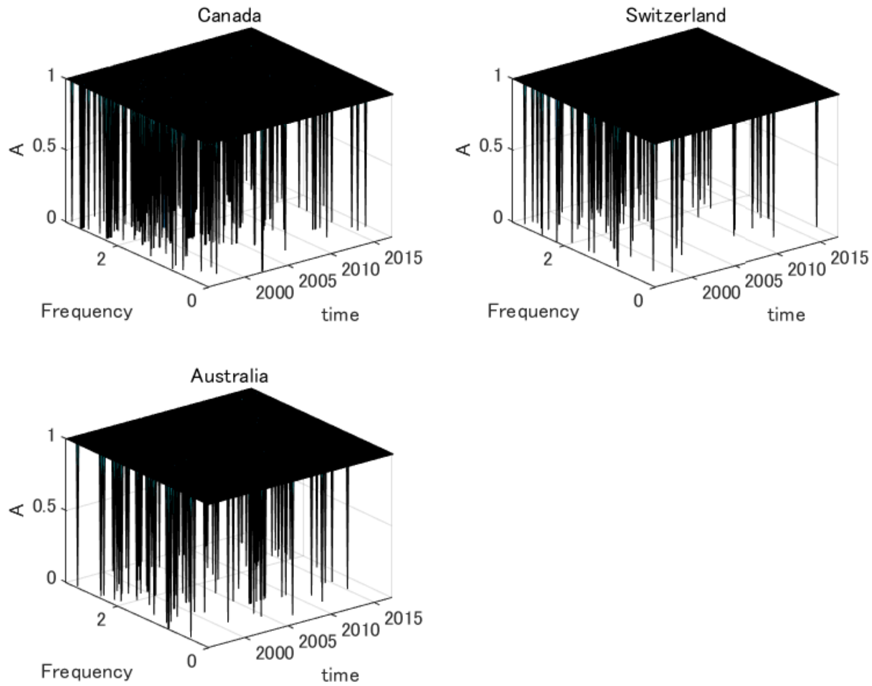

Figure 11(b): Estimated A for the PPP Fundamentals with $h = 12$.

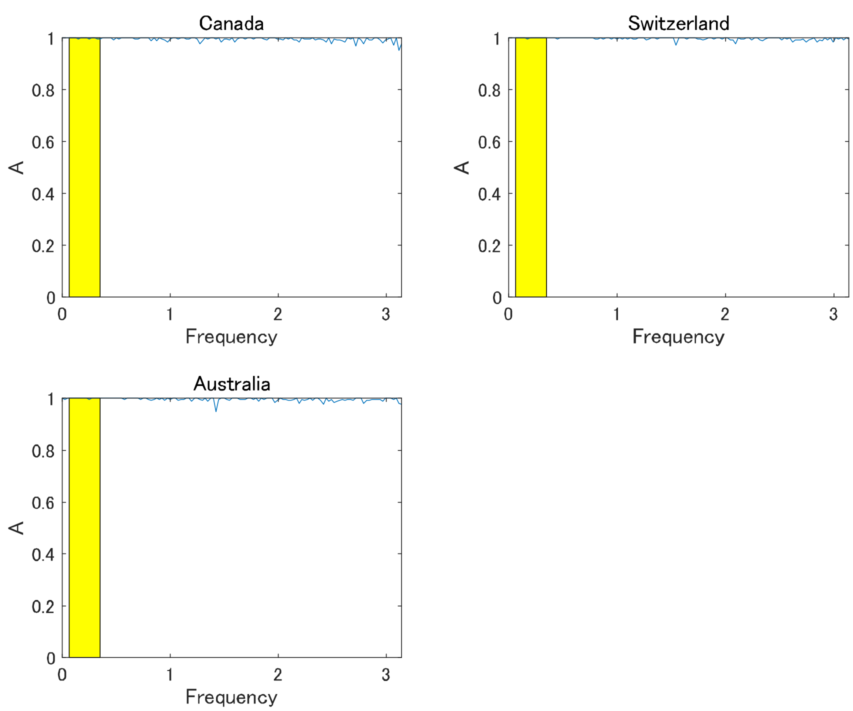

Figure 12(a): Estimated Average A for the PPP Fundamentals with $h = 24$.

(Business cycle frequencies are shaded.)

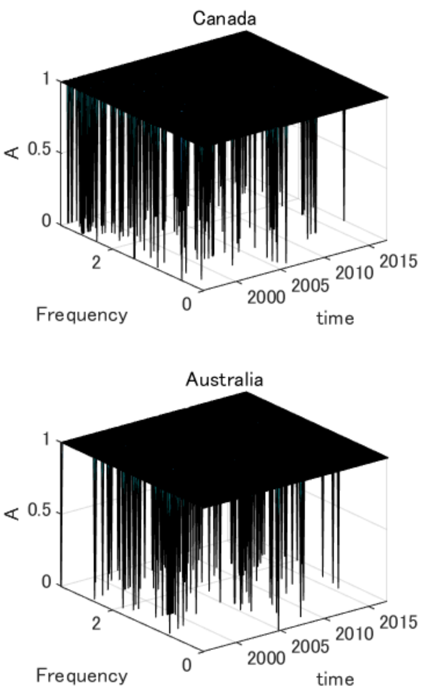

Figure 12(b): Estimated A for the PPP Fundamentals with $h = 24$.

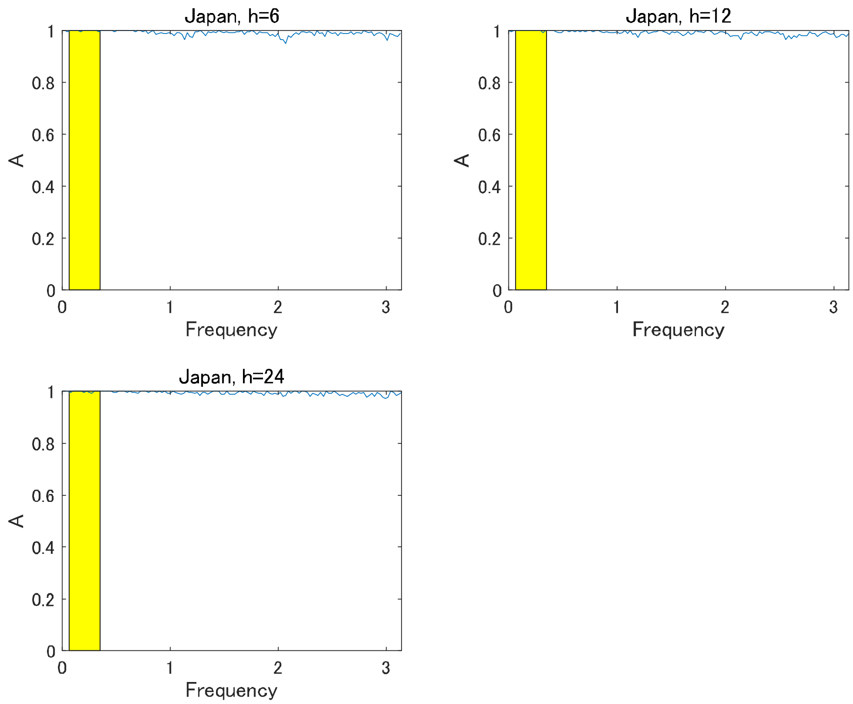

Figure 13(a): Estimated Average A for Monetary Fudamentals with $h = 6$, 12, and 24.

(Business cycle frequencies are shaded.)

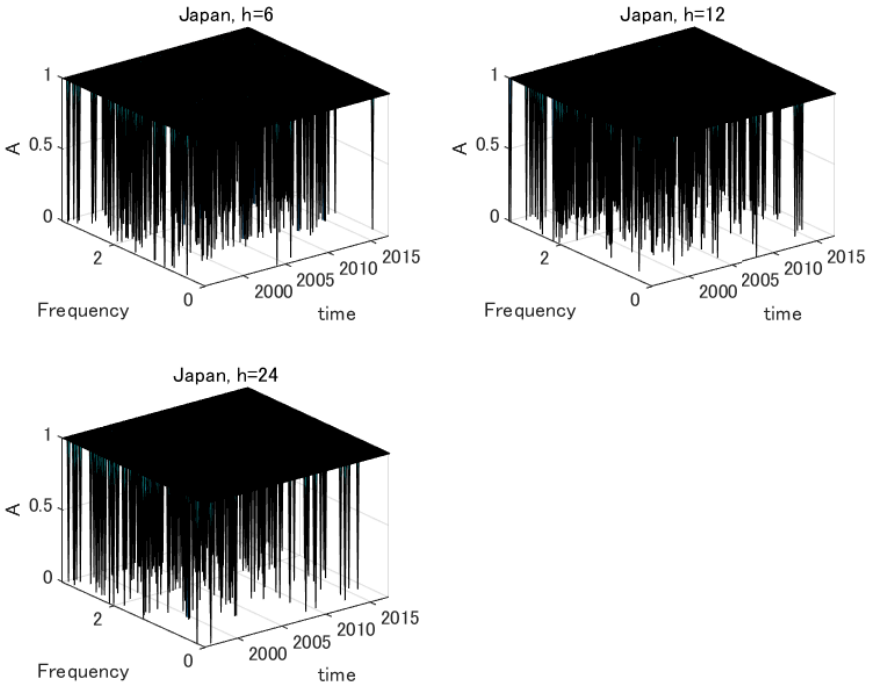

Figure 13(b): Estimated A for Monetary Fudamentals with $h = 6$, 12, and 24.

References

[1] Clark, T. E. and M.W. McCracken (2012) "Reality Checks and Comparisons of Nested Predictive Models," *Journal of Business and Economic Statistics* 30 (1), 53-66.

[2] Clark, T. E. and M.W. McCracken (2013) "Advances in Forecast Evaluation," in Elliott, G. and A. Timmermann Eds. *Handbook of Economic Forecasting*,Volume 2B.

[3] Corbae D., S. Ouliaris, and P.C.B. Phillips (2002) "Band Spectral Regression with Trending Data," *Econometrica* 70 (3), 1067-1109.

[4] Sargent, T. (1987) *Macroeconomic Theory*, 2nd Ed. Boston : Academic Press.

[5] Sims, C. (1972) "The Role of Approximate Prior Restrictions in Distributed Lag Estimation," *Journal of the American Statistical Association*, 67 (337), 169-175.