Online Appendix
International Stock Market Efficiency: A Non-Bayesian Time-Varying Model Approach

Mikio Ito[a], Akihiko Noda[b*] and Tatsuma Wada[c]

[a] *Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan*

[b] *Faculty of Economics, Wakayama University, 930 Sakaedani, Wakayama 640-8510, Japan*

[c] *Department of Economics, Wayne State University, Faculty and Administration Building, 656 W.Kirby St., Detroit, MI, 48202*

Abstract: This note provides detailed computations in the main text.

Keywords: Non-Bayesian Time-Varying VAR Model; International Linkages; Market Efficiency.

JEL Classification Numbers: C22; G14; G15.

A.1 Parameter Constancy Test for Time-Invariant VAR Model

There are some parameter constancy tests, for example Andrews (1993) and Nyblom (1989). Hansen (1992) develops a parameter constancy test for linear and non-linear models. The test is made under the null and alternative hypotheses: (H_0) the parameters are constant over time and (H_1) they follow a martingale process. In practice, we reformulate Equation (1) to extend Hansen's test for a VAR(p) model.

Define the $(k \times T)$ data matrix as

$$Y := (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_T),$$

where T is the number of time series observations. To reformulate the VAR(p) model with a intercept term into a regression formula, we introduce the following auxiliary vector and matrix:

$$Z_t := \begin{pmatrix} 1 \\ \boldsymbol{y}_t \\ \vdots \\ \boldsymbol{y}_{t-p} \end{pmatrix},$$

and

$$Z := (Z_0, Z_1, \cdots, Z_{T-1}).$$

Note that Z_0 is regarded as a prior of estimation. Defining the $(k \times kT)$ disturbance matrix as

$$U := (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_T),$$

*Corresponding Author. E-mail: noda@eco.wakayama-u.ac.jp, Tel. +81-73-457-7712, Fax. +81-73-457-7713.

we obtain a matrix form of Equation (1):

$$Y = BZ + U,$$

where a $k \times (1 + kp)$ matrix B is $[\boldsymbol{\nu}, A_1, \cdots, A_p]$. Using the *vec* operator, which transforms a $(k \times T)$ matrix into an (kT) vector by stacking the columns, we have the following regression form of Equation (1):

$$vec(Y) = (Z' \otimes I_k)vec(B) + vec(U), \tag{A.1}$$

where $\otimes$ is the Kronecker product and I_k denotes k identity matrix. In case $k = 2$ and $p = 1$, Equation (A.1) is specifically exhibited as

$$\begin{pmatrix} y_{11} \\ y_{21} \\ y_{12} \\ y_{22} \\ \vdots \\ y_{1T} \\ y_{2T} \end{pmatrix} = \begin{pmatrix} 1 & 0 & y_{10} & 0 & y_{20} & 0 \\ 0 & 1 & 0 & y_{10} & 0 & y_{20} \\ 1 & 0 & y_{11} & 0 & y_{21} & 0 \\ 0 & 1 & 0 & y_{11} & 0 & y_{21} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & y_{1,T-1} & 0 & y_{2,T-1} & 0 \\ 0 & 1 & 0 & y_{1,T-1} & 0 & y_{2,T-1} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \alpha_{11} \\ \alpha_{21} \\ \alpha_{12} \\ \alpha_{22} \end{pmatrix} + \begin{pmatrix} u_{11} \\ u_{21} \\ u_{12} \\ u_{22} \\ \vdots \\ u_{1T} \\ y_{uT} \end{pmatrix}.$$

In order to extend the procedure of Hansen (1992), we rewrite Equation (A.1) as

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u},$$

where $\boldsymbol{y} = vec(Y)$, $X = Z' \otimes I_k$ and $\boldsymbol{u} = vec(U)$. Given $\boldsymbol{y}$, we obtain the OLS estimate $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\boldsymbol{y}$ and the estimate of the covariance matrix $\hat{\Sigma}_u$ as UU'/T or $UU'/(T - kp - 1)$ (see Lütkepohl (2005, ch.3) for details).

Furthermore, to extend Hansen's (1992) procedure to our system, let us organize the above equation as follows:

$$\boldsymbol{y}_t = X_t\boldsymbol{\beta} + \boldsymbol{u}_t, \ t = 1, 2, \cdots, T. \tag{A.2}$$

For simplicity, we consider an OLS estimation of equation (A.2) with the i.i.d. assumption, $E[\boldsymbol{u}_t\boldsymbol{u}_t'] = \Sigma_u, t = 1, 2, \cdots$. The OLS estimates $\hat{\boldsymbol{\beta}}$ as well as $\hat{\Sigma}_u$ should hold the following equations.

$$\mathbf{0} = \sum_{\tau=1}^{T} X_\tau' \boldsymbol{e}_\tau$$

and

$$\mathbf{0} = \sum_{\tau=1}^{T} (vech(\boldsymbol{e}_\tau \boldsymbol{e}_\tau') - vech(\hat{\Sigma}_u)),$$

where $\boldsymbol{e}_t = \boldsymbol{y}_t - X_t\hat{\boldsymbol{\beta}}$ and the *vech* operator is closely related to *vec*, which only stacks the elements on and below the main diagonal of a symmetric matrix. To construct Hansen's (1992) test statistic for a simultaneous linear equation system, define $(k(1 + kp) + k(k + 1)/2)$ vector,

$$f_\tau(\hat{\boldsymbol{\beta}}, \hat{\Sigma}_u) = \begin{bmatrix} vec(X_\tau' \boldsymbol{e}_\tau) \\ vech(\boldsymbol{e}_\tau \boldsymbol{e}_\tau') - vech(\hat{\Sigma}_u) \end{bmatrix}, \ \tau = 1, 2, \cdots, T.$$

Let $f_{i\tau}$ denotes i-th component of $f_\tau(\hat{\boldsymbol{\beta}}, \hat{\Sigma}_u)$, $S_s = \sum_{\tau=1}^{s} f_\tau$ and $S_{is} = \sum_{\tau=1}^{s} f_{i\tau}$ for $i = 1, 2, \cdots, k(1+kp) + k(k+1)/2$, $s = 1, 2, \cdots, T$. Note that we abbreviate $\hat{\boldsymbol{\beta}}, \hat{\Sigma}_u$ when there is no confusion. Hansen's (1992) individual statistics of parameter constancy test are

$$L_i = \frac{1}{TV_i} \sum_{\tau=1}^{T} S_{is}^2, \ i = 1, 2, \cdots, (k(1+kp) + k(k+1)/2),$$

where $V_i = \sum_{\tau=1}^{T} f_{i\tau}^2$. His joint statistic is defined:

$$L_c = \frac{1}{T} \sum_{\tau=1}^{T} S'_\tau V^{-1} S_\tau,$$

where $V = \sum_{\tau=1}^{T} f_\tau f'_\tau$. These test statistics follow singular distributions represented by the Brownian motions and bridges. See Hansen (1990) for the tables of the statistics.

A.2 Regressor of State Space Regression Models

This section shows the reader a significant property of the method that we present for dealing with a TV-VAR(p) model in Section 3. The method enables us to handle the model as a conventional econometric model by identifying the corresponding state space model with a linear regression model. Recall that the observation equation and the state one appeared in the state space model in Section 3: Equations (3.3) and (3.4).

We prove that estimating method of the smoother of a state space model guarantees the exact estimate whatever data is applied. Mathematically we assert that the regressor matrix of our underlying linear system has full rank. This implies that our estimation never suffers from multicolinearity and that our model is always identifiable in the sense of Rothenberg (1971). The following results do not depend on whatever time-varying model is assumed; they cover our TV-VAR model, a typical state space model.

The model consists of two equations as follows:

$$\boldsymbol{y}_t = X_t \boldsymbol{\beta}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \overset{iid}{\sim} \mathcal{WS}(\boldsymbol{0}, \boldsymbol{\Sigma_u}), \quad t = 1, 2, \cdots, T, \tag{A.3}$$

and

$$\boldsymbol{\beta}_{t+1} = B_t \boldsymbol{\beta}_t + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \overset{iid}{\sim} \mathcal{WS}(\boldsymbol{0}, \boldsymbol{\Sigma_v}), \quad t = 1, 2, \cdots, T, \tag{A.4}$$

where $\mathcal{WS}(\boldsymbol{0}, \boldsymbol{\Sigma_u})$ and $\mathcal{WS}(\boldsymbol{0}, \boldsymbol{\Sigma_v})$ are any distributions depending only on their means and variances such as the normal distributions. Note that $\boldsymbol{y}_t$'s are $k-$dimensional vectors and $\boldsymbol{\beta}_t$'s are $m-$dimensional vectors as parameters. X_t and B_t are $k \times m$ and $m \times m$ matrices respectively. Notice that two disturbances, $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$, follow iid normal distributions with zero means and covariance matrices, $\boldsymbol{\Sigma_u}$ and $\boldsymbol{\Sigma_v}$, respectively.

By transforming Equation (A.4) to the following form, we can regard the state equation as a linear regression model.

$$\boldsymbol{0} = B_t \boldsymbol{\beta}_t - I \boldsymbol{\beta}_{t+1} + \boldsymbol{v}_t, \quad \boldsymbol{v}_t \overset{iid}{\sim} \mathcal{WS}(\boldsymbol{0}, \boldsymbol{\Sigma_v}), \quad t = 1, 2, \cdots, T. \tag{A.5}$$

Combining Equations (A.3) with (A.5) into the following linear system, we can employ conventional econometric techniques to obtain the Kalman smoother.

$$\hat{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X} \\ \bar{\boldsymbol{W}} \end{bmatrix} \boldsymbol{\beta} + \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}$$

where

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_T \end{pmatrix}, \quad \boldsymbol{\gamma} = \begin{pmatrix} -B_0\boldsymbol{\beta}_0 \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{u} = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_T \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_T \end{pmatrix},$$

and

$$\boldsymbol{X} = \begin{pmatrix} X_1 & & & \boldsymbol{O} \\ & X_2 & & \\ & & \ddots & \\ \boldsymbol{O} & & & X_T \end{pmatrix}, \quad \boldsymbol{W} = \begin{pmatrix} -I & & & \boldsymbol{O} \\ B_1 & -I & & \\ & \ddots & \ddots & \\ \boldsymbol{O} & & B_{T-1} & -I \end{pmatrix}.$$

Our main theorem of this appendix is as follows:

Theorem 1 *The regressor matrix has full rank. That is*

$$\text{rank} \begin{bmatrix} \boldsymbol{X} \\ \hdashline \boldsymbol{W} \end{bmatrix} = mT.$$

Proof: Let $\hat{\boldsymbol{X}}$ denote the regressor matrix. It is of size $(k+m)T \times mT$. Thus, $\text{rank}\hat{\boldsymbol{X}} \leq mT$. Clearly, each ranks of the two submatrices, $\boldsymbol{W}$ and $\boldsymbol{X}$, are fewer than or equal to mT. Formally,

$$\text{rank}\boldsymbol{X} \leq mT,$$

and

$$\text{rank}\boldsymbol{W} \leq mT.$$

On the other hand, $\text{rank}\boldsymbol{W} = \text{rank}\boldsymbol{W}'$. Since W' has the following reduced row echelon form:

$$\begin{pmatrix} -I & B_1 & & \boldsymbol{O} \\ & -I & \ddots & \\ & & \ddots & B_{T-1} \\ \boldsymbol{O} & & & -I \end{pmatrix},$$

its rank is clearly mT. Finally, $mT \leq \text{rank}\hat{\boldsymbol{X}} \leq mT$. This implies $\text{rank}\hat{\boldsymbol{X}} = mT$. $\square$

The reader should notice that the rank of $\hat{\boldsymbol{X}}$ depends neither on the data matrices, $X_1, X_2, \cdots, X_T$, nor on the state transition matrices, $B_1, B_2, \cdots, B_{T-1}$. This signifies that classical least square techniques such as OLS or GLS are always applicable to estimate linear time-varying models that can be represented as state space models.

In this paper, our TV-VAR(1) model with constant drift. Such a model has two types of parameters to be estimated: time-varying and time-invariant. In the rest of this appendix, we confirm that the above discussion holds even if linear time-varying model such as our model has time-invariant parameters. In the case, we should modify Equation (A.3) as follows:

$$\boldsymbol{y}_t = Z_t\boldsymbol{\alpha} + X_t\boldsymbol{\beta}_t + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \overset{iid}{\sim} \mathcal{WS}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{u}}), \quad t = 1, 2, \cdots, T,$$

where $\boldsymbol{\alpha}$ is a $\ell-$vector of time-invariant parameters and Z_t is a $k \times \ell$ matrix of data. The regressor matrix turns to be

$$\left[\begin{array}{cc} \boldsymbol{Z} & \boldsymbol{X} \\ \hdashline \boldsymbol{O} & \boldsymbol{W} \end{array}\right],$$

where the above $\boldsymbol{O}$ is $mT \times \ell$ matrix of zero and

$$\boldsymbol{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_T \end{pmatrix}.$$

The following theorem holds.

Theorem 2

$$\text{rank}\left[\begin{array}{cc} \boldsymbol{Z} & \boldsymbol{X} \\ \hdashline \boldsymbol{O} & \boldsymbol{W} \end{array}\right] = \text{rank}\boldsymbol{Z} + mT.$$

Proof: Let $\tilde{X}$ denote the regressor matrix in the case. According to Theorem 1, rank$\hat{\boldsymbol{X}}$ is mT. Thus, using the reduced row echelon form of $\boldsymbol{W}$, say, $\boldsymbol{W}_0$, we can transform $\hat{\boldsymbol{X}}$ to the following form:

$$\left[\begin{array}{c} \boldsymbol{O} \\ \hdashline \boldsymbol{W}_0 \end{array}\right].$$

Then there are two non-singular matrices, P of size $((k+m)T) \times ((k+m)T)$ and Q of size $(\ell + mT) \times (\ell + mT)$, such that

$$P\tilde{\boldsymbol{X}}Q = \left[\begin{array}{cc} \boldsymbol{Z} & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{W}_0 \end{array}\right].$$

Considering rank$\boldsymbol{Z} = mT$ and

$$\text{rank}\left[\begin{array}{c} \boldsymbol{Z} \\ \hdashline \boldsymbol{O} \end{array}\right] = \text{rank}\boldsymbol{Z},$$

we have rank$\tilde{\boldsymbol{X}} = \text{rank}\boldsymbol{Z} + mT$. $\square$

Note that we can deal with the case of the TV-VAR(1) with time-invariant drift term such as our model in this paper. As appeared in section 3.2, one regards $\boldsymbol{Z}$ as

$$\begin{pmatrix} I \\ I \\ \vdots \\ I \end{pmatrix},$$

where each I denotes $k \times k$ identity matrix. Then, rank$\boldsymbol{Z} = k$. Considering the number of parameters to be estimated is $k + k^2 T$ and $m = k^2$, our TV-VAR model has no problem of identification.

A.3 Non-Bayesian Time-Varying VAR Model

If the parameter constancy test exhibits instability of VAR coefficients over time, we choose an alternative model that holds the alternative hypothesis on the parameter dynamics and then estimate the model. Focusing our attention on linkage of stock markets which is supposed to vary over time, we suppose that only VAR coefficients vary over time while any components of the intercept term are invariant. In what follows, we present how to estimate the non-Bayesian TV-VAR model. Our approach has several good points in comparison with the Bayesian one (see, for example, Cogley and Sargent (2001, 2005) and Primiceri (2005)).

We start by setting VAR coefficients varying over time for an ordinary VAR(p) model as follows:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_{1,t}\boldsymbol{y}_{t-1} + \cdots + A_{p,t}\boldsymbol{y}_{t-p} + \boldsymbol{u}_t,\ t = 1, 2, \cdots, T.$$

This equation can be represented as follows:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + A_t Z_{t-1} + \boldsymbol{u}_t,\ t = 1, 2, \cdots, T, \tag{A.6}$$

where

$$A_t = [A_{1,t}\ \cdots\ A_{p,t}] \text{ and } Z_{t-1} = \begin{bmatrix} \boldsymbol{y}_{t-1} \\ \vdots \\ \boldsymbol{y}_{t-p} \end{bmatrix}.$$

Equation (A.6) corresponds to an observation equation when we regard our TV-VAR model as a state space model. As to the corresponding state equation, let us assume the following random walk process:

$$A_{i,t} = A_{i,t-1} + V_{i,t},\ i = 1, 2, \cdots, p \text{ and } t = 1, 2, \cdots, T,$$

where each $V_{i,t}$ is a $k \times k$ matrix of random variables, say, following normal distributions. Let V_t denote $= [V_{1,t}\ \cdots\ V_{p,t}]$ or equivalently

$$vec(A_t) = vec(A_{t-1}) + \boldsymbol{v}_t,\ t = 1, 2, \cdots, T, \tag{A.7}$$

where each $\boldsymbol{v}_t$ is a k^2p-vector of random variables, $vec(V_t) = vec([V_{1,t}\ \cdots\ V_{p,t}])$.

In place of the Kalman smoothing, we estimate $\boldsymbol{\nu}, A_1, \cdots, A_T$ by considering together Equations (A.6) and (A.7) as a simultaneous system of linear equations. Equation (A.6) turns out to be

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_T \end{pmatrix} = \begin{pmatrix} I_k & Z_0' \otimes I_k & & & \boldsymbol{O} \\ I_k & & Z_1' \otimes I_k & & \\ \vdots & & & \ddots & \\ I_k & \boldsymbol{O} & & & Z_{T-1}' \otimes I_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\nu} \\ vec(A_1) \\ vec(A_2) \\ \vdots \\ vec(A_T) \end{pmatrix} + \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_T \end{pmatrix}.$$

This is equivalent to the following equation.

$$\boldsymbol{y} = [\boldsymbol{1} \otimes I_k\ \ \text{diag}(Z_0' \otimes I_k, Z_1' \otimes I_k, \cdots, Z_{T-1}' \otimes I_k,)]\ vec([\boldsymbol{\nu}, A_1, A_2, \cdots, A_{T-1}]) + vec(U), \tag{A.8}$$

where $\mathbf{1} = (1\ 1\ \cdots\ 1)' \in \boldsymbol{R}^T$. On the other hand, the state equation can be represented as:

$$\begin{pmatrix} -vec(A_0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & -I_{k^2p} & & & \\ \mathbf{0} & I_{k^2p} & -I_{k^2p} & & \\ \vdots & & \ddots & \ddots & \\ \mathbf{0} & & & I_{k^2p} & -I_{k^2p} \end{pmatrix} \begin{pmatrix} \boldsymbol{\nu} \\ vec(A_1) \\ vec(A_2) \\ \vdots \\ vec(A_T) \end{pmatrix} + \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_T \end{pmatrix},$$

where $vec(A_0)$ is a prior vector of VAR coefficients, of whichever level is set, do not influence our results except starting burn-in periods.

Considering the following linear system, we estimate the coefficients of our time-varying VAR by using OLS.

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} D \\ W \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $D = \left[\mathbf{1} \otimes I_k \,|\, \mathrm{diag}(Z_0' \otimes I_k, Z_1' \otimes I_k, \cdots, Z_{T-1}' \otimes I_k)\right]$,

$$\boldsymbol{\gamma} = \begin{bmatrix} -vec(A_0) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix},$$

$$W = \begin{pmatrix} \mathbf{0} & -I_{k^2p} & & & \\ \mathbf{0} & I_{k^2p} & -I_{k^2p} & & \\ \vdots & & \ddots & \ddots & \\ \mathbf{0} & & & I_{k^2p} & -I_{k^2p} \end{pmatrix},$$

$$\boldsymbol{\beta} = vec([\boldsymbol{\nu}, vec(A_1), \cdots, vec(A_T)])$$

and

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}.$$

The OLS estimate is:

$$\hat{\boldsymbol{\beta}} = \left(\begin{bmatrix} D \\ W \end{bmatrix}' \begin{bmatrix} D \\ W \end{bmatrix} \right)^{-1} \begin{bmatrix} D \\ W \end{bmatrix}' \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{\gamma} \end{bmatrix}.$$

As Ito et al. (2016) show, the OLS estimate, $\hat{\boldsymbol{\beta}}$, of the above linear model provides the same estimation, $\hat{A}_t$, $(t = 1, 2, \cdots, T)$ as those of the Kalman smoothing for the original state space model. Not using iterative procedure of the traditional Kalman smoothing, our approach allows us to use familiar econometric techniques such as heteroskedastic autoeregressive consistent (HAC) estimations (see, for example, Newey and West (1987, 1994)). Such a HAC estimate enables us to obtain the covariance estimate of $\Sigma_{vec(A_t)}, (t = 1, 2, \cdots, T)$ from the covariance estimate, $\Sigma_{\boldsymbol{\beta}}$. It is possible to construct the confidence intervals of the VAR coefficients for period by period by using the squared diagonal components of $\Sigma_{\hat{\boldsymbol{\beta}}}$. Furthermore, as shown in Technical Appendix A.2, our approach never confronts with the problem of identifiability unlike usual econometric analyses using least square techniques. In comparison with Bayesian approaches, our method is so simple that we are free from sometimes agonizing choice of prior distribution of parameters (for a typical example of the Bayesian TV-VAR model approach, see Cogley and Sargent (2001, 2005) and Primiceri (2005)).

A.4 Monte Carlo Method for TV-VAR Estimations

This subsection provides our method of statistical inferences on the TV-VAR estimates and their derived statistics, the degree of market efficiency. The idea is so simple that the Monte Carlo technique brings about their confidence bands by under the hypothesis that any markets are efficient at any periods.

The practical procedure is as follows. We first estimate the means and standard deviations of time series of stock market returns over the periods, $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1 \cdots \hat{\mu}_k)'$ and $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_1 \cdots \hat{\sigma}_k)'$ using the original data. Then, we derive N time series samples with length T by a Monte Carlo method from an i.i.d. normal distribution with the following means and variance structure:

$$\boldsymbol{y}_t^{(n)} = \boldsymbol{u}_t^{(n)}, \quad \boldsymbol{u}_t^{(n)} \overset{iid}{\sim} \mathcal{WS}\left(\begin{pmatrix} \hat{\mu}_1 \\ \vdots \\ \hat{\mu}_k \end{pmatrix}, \begin{pmatrix} \hat{\sigma}_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{\sigma}_k^2 \end{pmatrix} \right), \ (n = 1, \cdots, N, \ t = 1, \cdots, T).$$

Note that we consdier as the efficient market hypothesis in the semi-strong sense as the null of our infernece. Thus, we suppose that the above uncorrelated disturbance vector $\boldsymbol{u}_t^{(n)}$, follows the above normal distribution. In other words, they are generated by a VAR process with time-invariant zero VAR coefficients.

Secondly we estimate the TV-VAR coefficients and residuals for our system of Equations (A.7) and (A.8) in section 3.2 from the above N artificial samples:

$$S_M := \{(\hat{A}_{1,t}^{(1)}, \cdots, \hat{A}_{p,t}^{(1)}, \hat{\boldsymbol{u}}_t^{(1)}, \hat{\boldsymbol{v}}_t^{(1)}), \cdots, (\hat{A}_{1,t}^{(N)}, \cdots, \hat{A}_{p,t}^{(N)}, \hat{\boldsymbol{u}}_t^{(N)}, \hat{\boldsymbol{v}}_t^{(N)})\}.$$

We derive N Monte Carlo samples of TV-VAR(p) model from S_M; it is natural to consider a distribution of the statistics with respect to each Monte Carlo samples. A set of N samples is available through the way shown in section 3.3 using the set of samples, S_M. Finally, we construct confidence bands from the N Monte Carlo samples.[1]

A.5 Bootstrap Method for TV-VAR Estimations

We can adopt another simulation method, bootstrap one, to attain the same goal as the Monte Carlo method does. The idea itself is very similar to that of the Monte Carlo technique in the previous section; it only differs by resampling process from the Monte Carlo technique.

The practical procedure is as follows. First we identify the stock returns data $\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_T\}$ with the residuals $D^0 = \{\hat{\boldsymbol{u}}_1, \cdots, \hat{\boldsymbol{u}}_T\}$ of VAR estimation under the hypothesis of all zero coefficients. Then we extract N samples $\{\tilde{\boldsymbol{y}}_1^{(i)}, \cdots, \tilde{\boldsymbol{y}}_T^{(i)}\}, i = 1, 2, \cdots, N$ with replace from D^0 regarding it as an empirical distribution of the residuals.

Secondly we estimate the TV-VAR coefficients and residuals for our system of Equations (A.7) and (A.8) in section 3.2 from the above N artificial samples:

$$S_b := \{(\tilde{A}_{1,t}^{(1)}, \cdots, \tilde{A}_{p,t}^{(1)}, \tilde{\boldsymbol{u}}_t^{(1)}, \tilde{\boldsymbol{v}}_t^{(1)}), \cdots, (\tilde{A}_{1,t}^{(N)}, \cdots, \tilde{A}_{p,t}^{(N)}, \tilde{\boldsymbol{u}}_t^{(N)}, \tilde{\boldsymbol{v}}_t^{(N)})\}.$$

We derive N bootstrap samples of TV-VAR(p) model from S_b; it is natural to consider a distribution of the statistics with respect to each bootstrap samples. A set of N samples is available through the way shown in section 3.3 using the set of samples, S_b. Finally, we construct confidence bands from the N bootstrap samples in the same way as the Monte Carlo technique.[2]

[1] Our Monte Carlo method is applicable to an AR(p) model for univarite data.

[2] Our bootstrap method is also applicable to an AR(p) model for univarite data.

References

Andrews, D. W. K. (1993), "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Cogley, T. and Sargent, T. J. (2001), "Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual*, 16, 331–373.

— (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US," *Review of Economic Dynamics*, 8, 262–302.

Hansen, B. E. (1990), "Lagrange Multiplier Tests for Parameter Instability in Non-linear Models," University of Rochester.

— (1992), "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling*, 14, 517–533.

Ito, M., Noda, A., and Wada, T. (2016), "The Evolution of Stock Market Efficiency in the US: A Non-Bayesian Time-Varying Model Approach," *Applied Economics*, 48, 621–635.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer.

Newey, W. K. and West, K. D. (1987), "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

— (1994), "Automatic lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223–230.

Primiceri, G. E. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

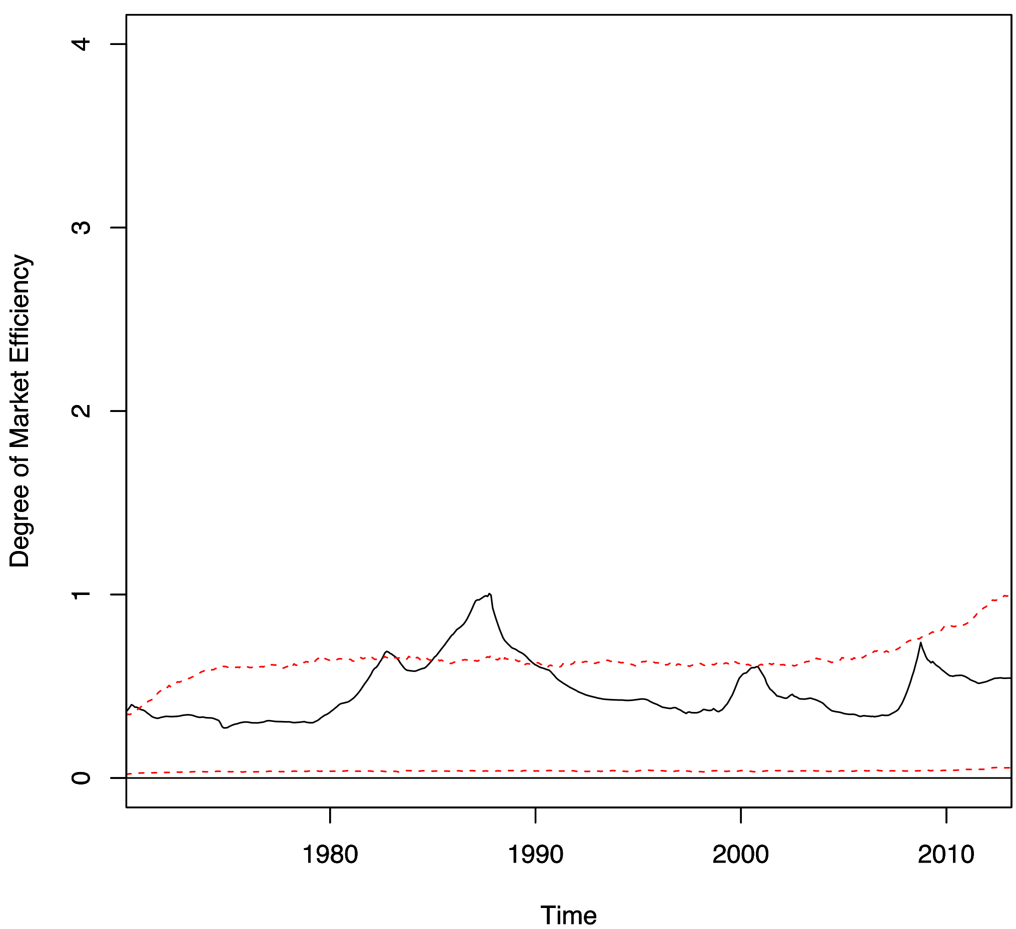
Rothenberg, T. J. (1971), "Identification in Parametric Models," *Econometrica*, 39, 577–591.

Table A.1: Time Invariant AR Estimations

	$Constant$	R_{t-1}	R_{t-2}	$\bar{R}^2$	L_C
R_t^{US}	0.0043	0.1975	$-$	0.0356	28.3517
	[0.0018]	[0.0528]			
R_t^{CA}	0.0041	0.2361	$-$	0.0521	41.0607
	[0.0020]	[0.0558]			
R_t^{GB}	0.0047	0.3172	-0.1222	0.0886	39.3477
	[0.0021]	[0.0628]	[0.0526]		
R_t^{JP}	0.0027	0.2738	$-$	0.0711	37.5094
	[0.0021]	[0.0453]			
R_t^{DE}	0.0030	0.2512	$-$	0.0596	27.2011
	[0.0022]	[0.0454]			
R_t^{FR}	0.0038	0.2299	$-$	0.0492	47.9710
	[0.0023]	[0.0442]			
R_t^{IT}	0.0026	0.2119	$-$	0.0412	44.5418
	[0.0027]	[0.0496]			

Note: As for table 2.

Figure A.1: Time-Varying Degree of Market Efficiency: North America



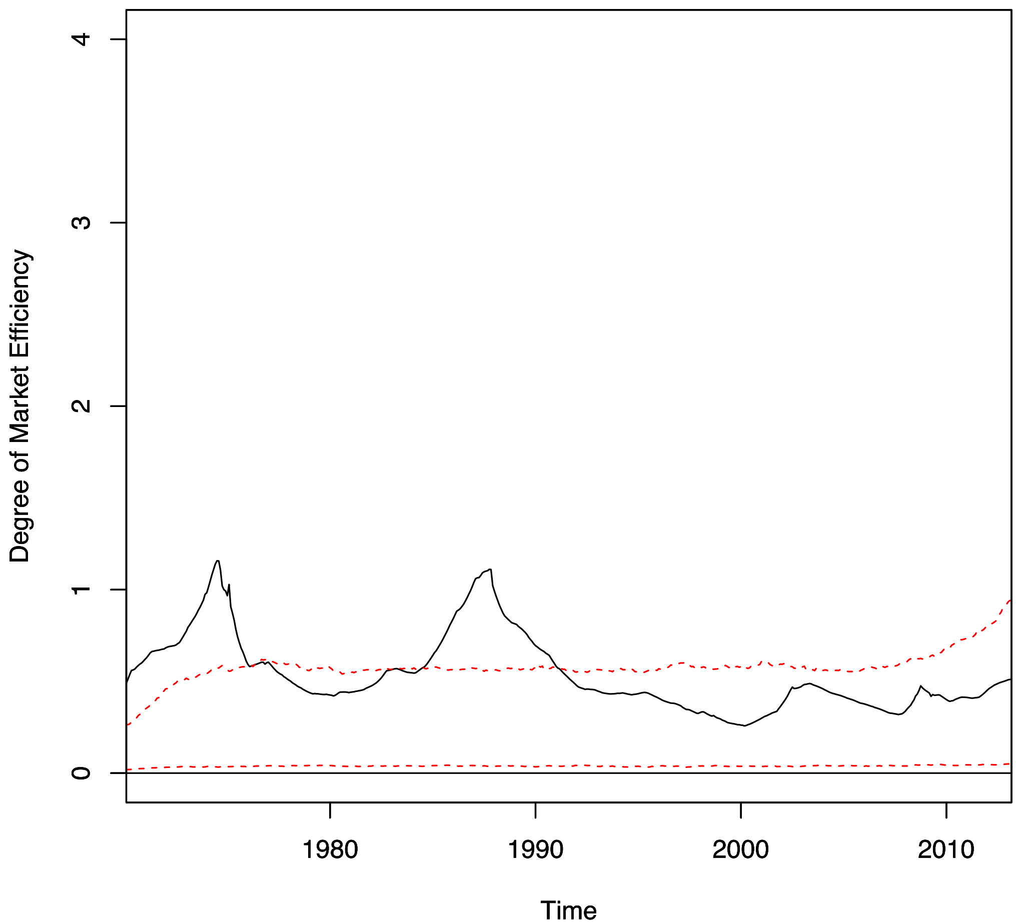
Notes:

(1) The dashed red lines represent the 99% confidence bands of the time-varying spectral norm in case of efficient market.

(2) We run 5000 times bootstrap sampling to calculate the confidence bands.

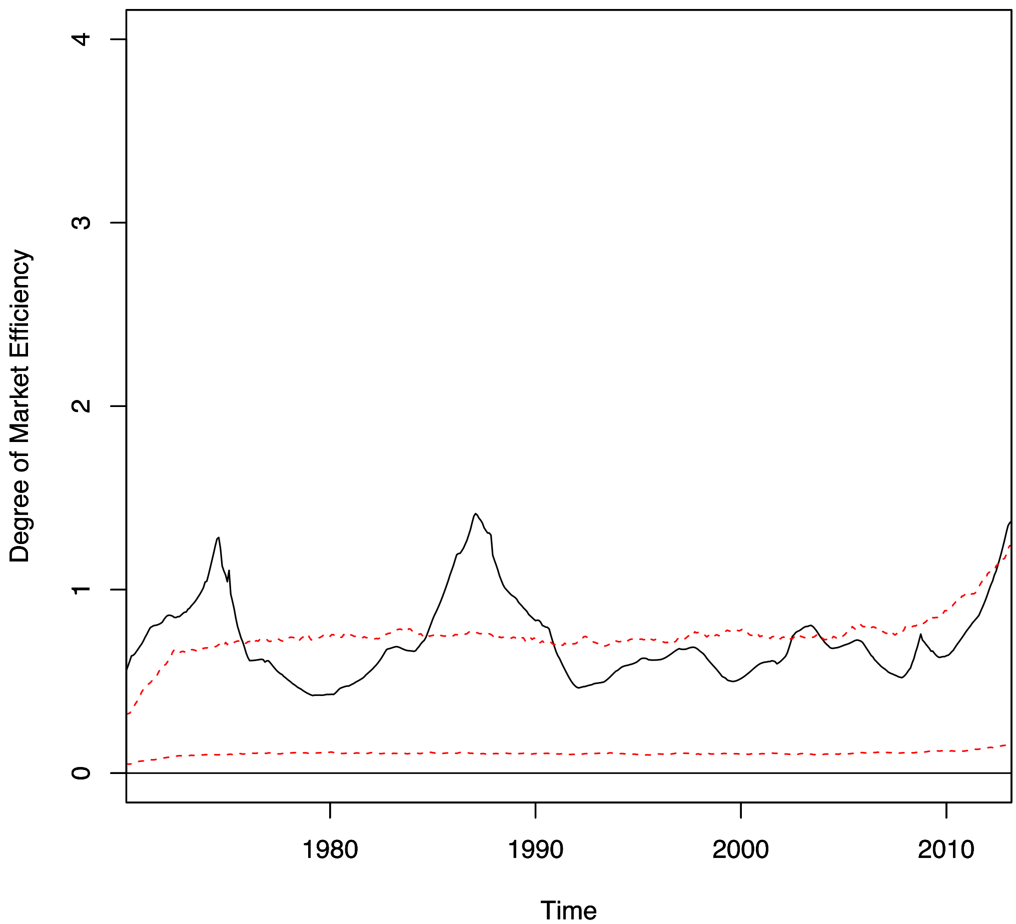
(3) R version 3.1.0 was used to compute the estimates.

Figure A.2: Time-Varying Degree of Market Efficiency: U.S. and U.K.



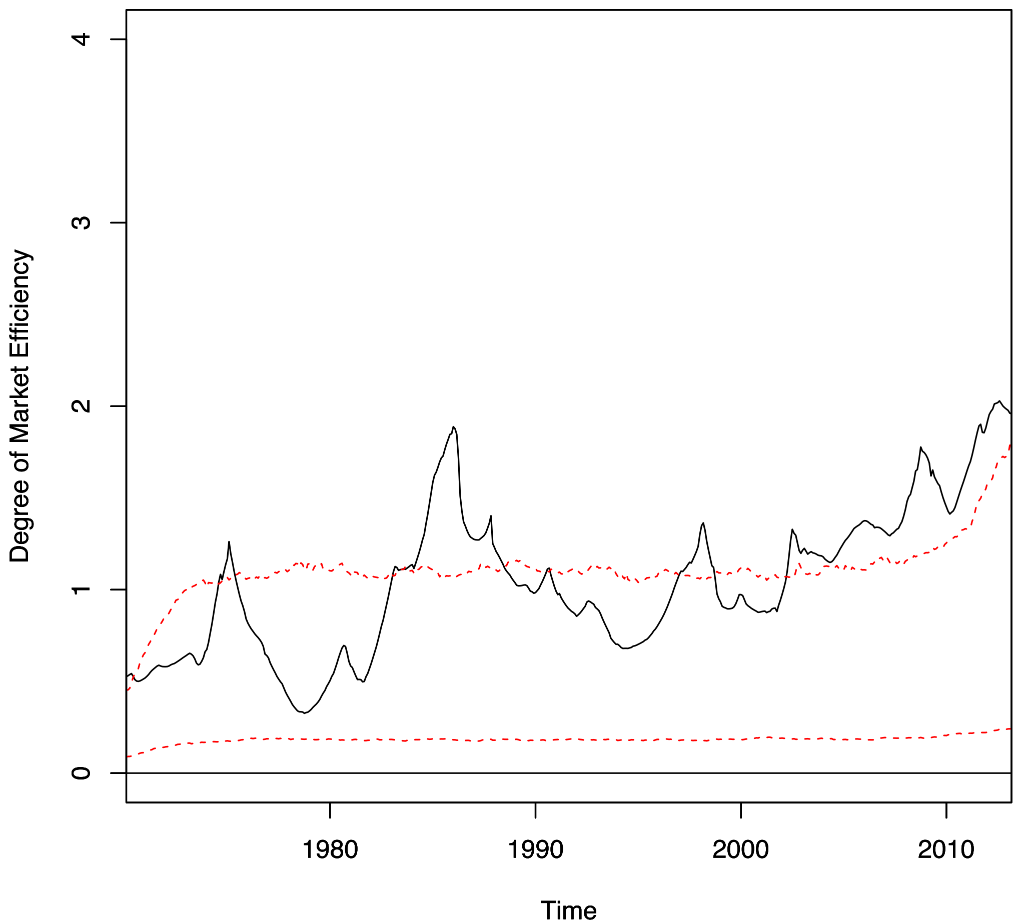
Note: As for figure A.1.

Figure A.3: Time-Varying Degree of Market Efficiency: U.S., U.K. and Japan



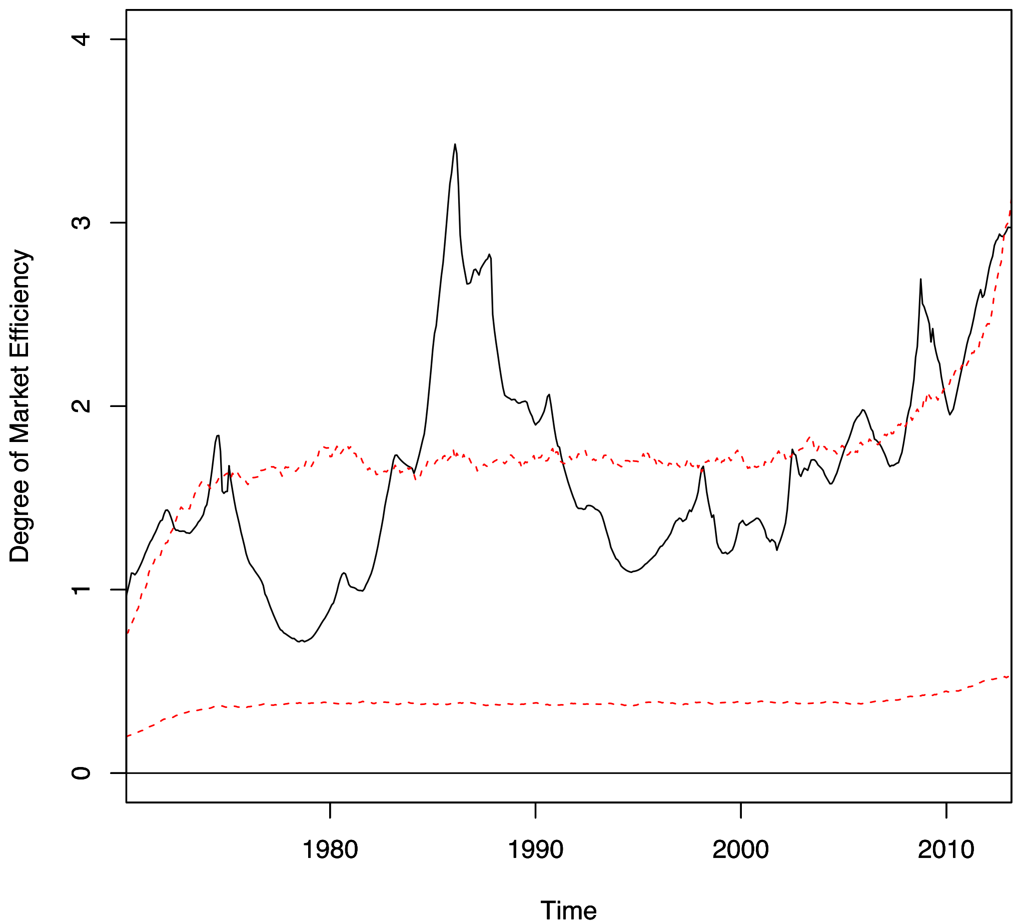
Note: As for figure A.1.

Figure A.4: Time-Varying Degree of Market Efficiency: European Countries



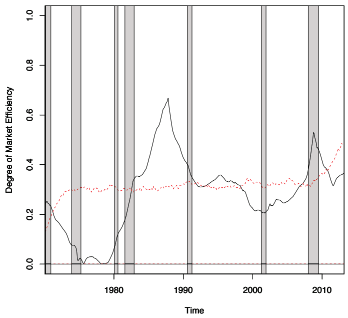
Note: As for figure A.1.

Figure A.5: Time-Varying Degree of Market Efficiency: G7 Countries

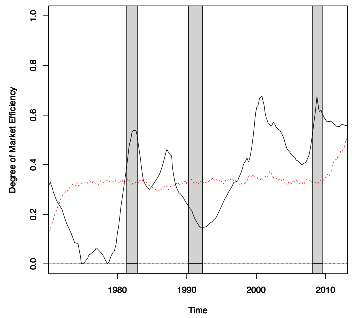


Note: As for figure A.1.

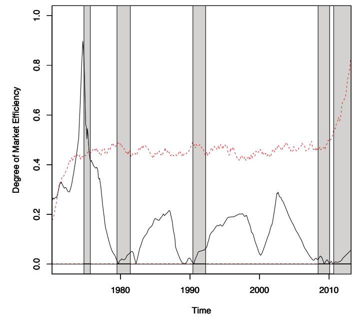
Figure A.6: Time-Varying Degree of Market Efficiency: Individual Countries



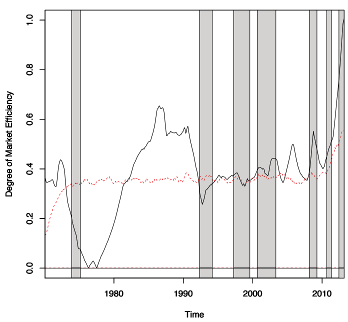
(a) U.S.



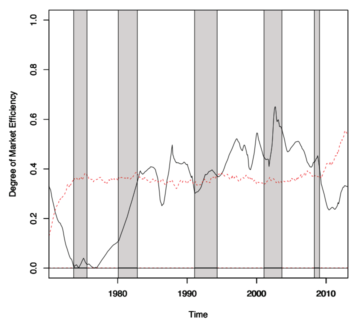
(b) Canada



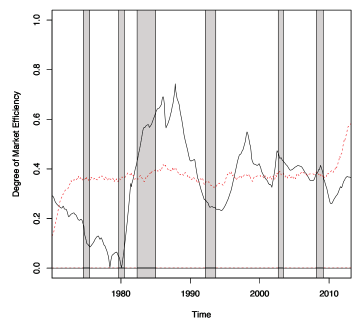
(c) U.K.



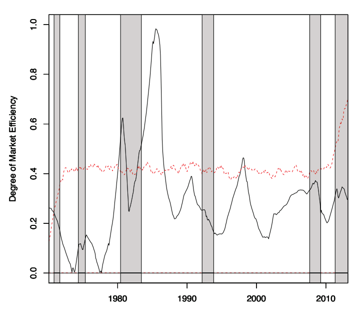
(d) Japan



(e) Germany



(f) France



(g) Italy

Notes:

(1) The dashed red lines represent the 99% confidence bands of the time-varying spectral norm in case of efficient market.

(2) We run 5000 times bootstrap sampling to calculate the confidence bands.

(3) The shade areas represent recessions reported by the NBER business cycle dates for the U.S. and the Economic Cycle Research Institute business cycle peak and trough dates for the other countries.

(4) R version 3.1.0 was used to compute the estimates.