Causality Detection in Financial Markets: Evolutionary Kernel-based Subset Time-Series

T.J. Brailsford
UQ Business School
University of Queensland, Australia

T.J. O'Neill
School of Finance and Applied Statistics
The Australian National University, Canberra, Australia

J. Penm*
School of Finance and Applied Statistics
The Australian National University, Canberra, Australia

Keywords: Evolutionary algorithms; Causality detection; Subset time-series.
JEL Classification: C10, C63, F31, G15

Abstract

In this paper we develop an evolutionary kernel-based time update algorithm to recursively estimate subset discrete lag models (including full-order models) with a forgetting factor and a constant term, using the exact-windowed case. For the first time ever in such models, the proposed recursions cover subset discrete lag models (SDL) with a forgetting factor. The algorithm applies to causality detection when the true relationship occurs with a continuous or a random delay. We then present two illustrations to demonstrate the use of the proposed evolutionary algorithm. In the first illustration we apply the proposed estimation algorithm to investigate the relationship between Australian's ten-year Commonwealth Treasury bond futures contracts and the underlying bond assets. The findings confirm the existence of instantaneous causal and bi-directional feedback relationships between this bond and futures markets. In the second illustration we apply the algorithm to identify the relationship between Australian's All Ordinaries Share Price Index and Australian's All Ordinaries Share Price Index futures. The findings also show the existence of instantaneous causal and bidirectional feedback relationships between the Australian stock and index futures markets.

* Contact author: Dr Jack Penm, School of Finance and Applied Statistics, Faculty of Economics and Commerce, The Australian National University, ACT 0200, Australia. Email: jack.penm@anu.edu.au.

1. Introduction

Powerful computing equipment has had a dramatic impact on financial time series analysis, and has motivated development of innovative approaches to financial time series modelling and data processing (Refenes *et al*, 1996). New methodology, involving the evolutionary kernel-based recursive subset time series approach, specifies models in a more sophisticated learning manner and uses the data in highly adaptive ways. Against that background this paper applies an innovative and evolutionary kernel smoothing algorithm to financial data in new and important areas, exemplified in the development and application of novel sequential recursive estimation of subset discrete lag (SDL) models using the exact-windowed case.

SDL models, including full-order models as a special case, are often necessary. Holmes and Hutter (1989) apply the SDL modelling which involves a dependent variable y(t) and a set of the current and/or lagged x(t) where there is a continuous or random delay. They also present the case that the Granger causal relationship between money supply and income in the U.S. is weak and may occur with a delay. Further, subset time-series models are often relevant, especially when measurements exhibit some form of periodic behaviour with a range of different natural periods. In particular, a sampled discrete transfer function includes the absent lagged variables in both its denominator and numerator, and this transfer function represents the structure of a underlying system with a subset structure. Most important, if the underlying true SDL process has a subset structure, the suboptimal model specification (for instance, a full-order structure) can give rise to inefficient estimates and inferior projections. When interpreting research findings it is impractical to ignore the possibility of zero coefficients in SDL models, particularly when periodic responses are likely.

Also, financial time series model builders are often concerned that the coefficients of their established models may not be constant over time, but vary when the models are disturbed by changes arising from outside factors. This has motivated researchers to develop sequential estimation algorithms that allow the coefficients to slowly evolve. In SDL modelling, while there are well developed sequential fitting algorithms for the full-order models (Kalouptsidis *et al*, 1984), these algorithms are not applicable to SDL models with a subset structure. This is because the 'optimal' subset model at time instant t may become 'suboptimal' at future time points. If one simply sets zero values for the coefficients of the missing lags and then applies the recursive algorithms for the full-order case, this will lead to a loss of efficiency in the model performance, as the subset structure of the model is not updated properly.

Further, in order to make effective use of parallel processing equipment and thereby gain computational speed, recursive algorithms are to be developed. These algorithms are computationally efficient; avoid cumbersome matrix inversion; and provide the obvious relations to update subset models at consecutive time instants. There will be practical applications in complexity management, developed for use in financial markets through the use of new SDL recursive algorithms.

Recently the application of the forgetting factor approach (Brailsford *et al*, 2005) to financial time-series analysis, in a subset autoregressive (AR) framework, has become widespread. Specifically, the forgetting factor has been widely used to capture non-stationarity through a slowly time-varying AR model. A subset AR model, which works well in explaining the behaviour of a process over a small sample in a given time period, may have to be augmented for a longer data span which evolves slowly over time due to economic, political or structural changes. Consequently, the forecasts obtained by allocating greater weight to more recent observations and 'forgetting' some of the past are likely to outperform alternatives in which such an allocation is not adopted. Forgetting factors can be both fixed and variable. Gijbels *et al* (1999), propose a relationship between a fixed forgetting factor λ and the bandwidth h of a kernel in kernel smoothing, where $h = (x_T - x_1)/-(T-1)\log_e \lambda$, with a time-series observed at equal-spaced time points $x_1, \ldots, x_T$. Brailsford *et al* (2005) derive a $h_{T-i} = (x_T - x_{T-i}) / -\log_e \lambda(T-i)$ relationship between a variable forgetting factor at time t, $\lambda(t)$, and the bandwidth of a kernel, and thus establish the linkage between the variable forgetting factor approach and kernel smoothing. SDL models using the forgetting factor have not been applied, however, to a wide range of problems arising in financial time-series modelling, estimation and simulations. This motivates development of novel evolutionary kernel-based recursions for SDL models. In this paper we focus on the fixed forgetting factor and use the exact-windowed case for data to undertake modelling investigation.

The remainder of the paper is organised as follows. In Section two we present the algorithm for recursively estimating SDL models. The proposed recursions cover, for the first time, SDL models, with a forgetting factor and a constant term. It is noteworthy that these recursions can be applied to vector SDL models (including both subset and full-order cases) in a straightforward fashion. In Section three it is shown that the current forgetting factor inclusive algorithm using the exact-windowed case is quite different from the algorithm, using the pre-windowed case. In Section four an illustration of the proposed evolutionary algorithm is used to describe the provision of all possible model structures for SDL modelling. In Section five, two illustrations are presented to demonstrate the practical use of the algorithm. The first application concerns a causality relationship between the bond and futures markets. The second application concerns the relationship between an Australian share-futures market. In Section six, a summary is given.

2. New evolutionary recursive algorithms for SDL modelling

In this section we introduce forward time-update recursions which recursively estimate a SDL model for the exact-windowed case.

In SDL modelling, it is desirable to relate $y(t)$ to present and past data for $x(t)$. We consider a SDL model of the form

$$y(t)+\rho+\sum_{i=1}^{p} a_i(I_s)x(t+1-i)=\varepsilon^a(t) \quad \text{and} \quad \{a_i(I_s)=0, \quad \text{as } i\in I_s,\}, \tag{2.1}$$

where ρ is a constant term, the order of the system is p, $1\le i_1 < i_2 \cdots < i_s \le p-1$; $x(t+1-i)$ is a deleted lag $(i\in I_s)$; a_i, i=1,2,...,p are parameters, and $\varepsilon^a(t)$ is a stationary process with E{ $\varepsilon^a(t)$ }=0 and

$$E\{\varepsilon^a(t)\varepsilon^a(t-\tau)\}=\begin{cases}\Omega & \tau=0\\ 0 & \tau\neq 0\end{cases}.$$

Equation (2.1) and properties associated with $\varepsilon^a(t)$ together constitute SDL, which involves a regressand $y(t)$ and a regressor $x(t)$. The set I_s specifies the integers between 1 and p-1 that correspond to excluded parameters.

Given two finite data sample sets, {x(n),..., x(T)} and {y(n),..., y(T)}, it is necessary to sequentially estimate all possible SDL models from (2.1) using the exact-windowed case. Since the actual scheme of (2.1) may not be order p, the resulting estimates of a_i is denoted by $a_{p,n,T}(i)$, where T is the sample size under examination. Then the predictor of a SDL system of (2.1) can be described as

$$\hat{y}(i)=-A'_{p,n,T}(I_s)X_{p-1,i}(I_s) \quad , \tag{2.2}$$

where $A'_{p,n,T}=\left[a_{p,n,T}(1),\rho_{p,n,T},a_{p,n,T}(2),\ldots,a_{p,n,T}(p)\right]$, and

$X_{p,i}=\left[x(i),\quad 1,\quad \ldots,\quad x(i-p)\right]'$.

$A_{p,n,T}(I_s)$ is formed by removing $a_{p,n,T}(i_1),\ldots,a_{p,n,T}(i_s)$ of $A_{p,n,T}$, and $X_{p,i}(I_s)$ is formed by removing $x(i+1-i_1),\ldots,x(i+1-i_s)$ of $X_{p,i}$.

The residual for observation i is

$$\eta_{p,n,i}(I_s) = y(i) + A'_{p,n,T-1}(I_s)X_{p-1,i}(I_s) \quad .$$

In reality, many time-series systems present complex non-stationary features and cannot be modelled by assuming that y(t) and x(t) are stationary. Thus, an estimate of the structure at time t should give a higher weight to the more recent observations and a lower weight to the observations of the more distant past. Thus, for a SDL model fitted by these two sample sets, we have

$$R_{p-1,n,T}(I_s)A_{p,n,T}(I_s) = -r_{p-1,n,T}(I_s) \quad \text{where } R_{p-1,n,T}(I_s) = \sum_{i=p-1+n}^{T} \lambda^{T-i} X_{p-1,i}(I_s)X'_{p-1,i}(I_s)$$

$$r_{p-1,n,T}(L_s) = \sum_{i=p-1+n}^{T} \lambda^{T-i} X_{p-1,i}(I_s)y(i), \quad \text{and } \Omega_{p,T}(I_s) = \sum_{i=p-1+n}^{T} \lambda^{T-i} \eta^2_{p,n,i}(I_s),$$

where λ, $0 < \lambda \le 1$, is the fixed forgetting factor as described in Hannan and Deistler (1988).

To develop time update recursions for SDL modelling, we consider the forward AR (p, I_s) model with a constant term of the form

$$x(t) + \tau + \sum_{i=1}^{p} h_i(I_s)x(t-i) = \varepsilon(t), \quad h_i(I_s) = 0, \text{ as } i \in I_s, \tag{2.3}$$

where ε(t) is an independent and identically distributed random process with

$$E\{\varepsilon(t)\} = 0, \; E\{\varepsilon(t)\,\varepsilon(t-k)\} \quad \begin{aligned} &= U(I_s) \quad \text{as} \quad k = 0 \\ &= 0 \qquad\;\; \text{as} \quad k \ne 0 \end{aligned} .$$

We also consider a backward AR(q) model of the form

$$g_p(M_s)x(t)+\beta+\sum_{i=0}^{p-1} g_i(M_s)x(t-p+i)=\bar{\varepsilon}(t),$$

$$g_0(M_s)=1,\ g_i(M_s)=0,\ \text{as } i\in M_s, \tag{2.4}$$

where $E\{\bar{\varepsilon}(t)\}=0,$ and the disturbance variance is $\bar{U}(M_s)$, M_s represents an integer set with elements m_1, m_2, ..., m_s, m_j=p-i_j, j = 1,2,...,s. A reciprocal integer pair for a forward subset AR model and a backward subset AR model is a pair of (2.3) and (2.4). Figure 1 shows a lag tree diagram which illustrates the reciprocal integer pairs of all subset AR processes up to and including lag length, $k=4$. Note that numerals represent particular lags in a forward AR and underlined numerals represent such leads in a backward AR.

We need to sequentially estimate all possible subset AR models from (2.3) and (2.4) using the exact-windowed case. Then we define observation i

$$\underline{X}'_{p,i}=[\,1,x(i-1),..,x(i-p)], \tag{2.5}$$

$$H'_{p,n,T}=[\tau_{p,n,T_p},h_{p,n,T}(1),...,h_{p,n,T}(p)],\ G'_{p,n,T}=[\,g_{p,n,T}(p),\xi_{p,n,T},...,g_{p,n,T}(1)]\ ,$$

For a reciprocal integer pair of the forward AR(p,I_s) and the backward AR(p,M_s) models fitted to this sample set, we have

$$R_{p,n,T}(I_s)=\sum_{i=p+n}^{T}\lambda^{T-i}X_{p,i}(O_s)X'_{p,i}(O_s),\ \text{where } X_{p,i}(O_s)=\begin{bmatrix} x(i) \\ \underline{X}_{p,i}(L_s)\end{bmatrix}=\begin{bmatrix} X_{p-1,i}(O_s) \\ x(i-p)\end{bmatrix},$$

$$R_{p,n,T}(I_s)\begin{bmatrix} 1 \\ \hline H_{p,n,T}(I_s)\end{bmatrix} = \begin{bmatrix} U_{p,n,T}(I_s) \\ \hline 0 \end{bmatrix},\ U_{p,n,T}(I_s) = \sum_{t=p+n}^{T} \lambda^{T-i}\left[\varepsilon_p(t)\right]^2 ,$$

where O_s represents an integer set with elements o_j, j=1,...,s, and $o_j = i_j + 2$. $X_{p,i}(O_s)$ and $X_{p-1,i}(O_s)$ are formed by removing the (o_1), ..., (o_s)'th row of $X_{p,i}$ and $X_{p-1,i}$ respectively. L_s represents an integer set with elements l_j, and $l_j = i_j + 1$. and $\underline{X}_{p,i}(L_s)$ is formed by removing the (l_1), ..., (l_s)'th row of $\underline{X}_{p,i}$. Also $H_{p,n,T}(I_s)$ is formed by removing $h_{p,n,T}(i_1)$, ..., $h_{p,n,T}(i_s)$ of $H_{p,n,T}$.

Now we define

$$C_{p,n,T}(I_s) = \sum_{i=p+n}^{T} \lambda^{T+1-i} \underline{X}_{p,i}(L_s)\underline{X}'_{p,i}(L_s), \qquad K_{p,n,T}(I_s) = C^{-1}_{p,n,T}(I_s)\underline{X}_{p,T+1}(L_s), \qquad \text{and}$$

$$\tau_{p,n,T}(I_s) = 1 + \underline{X}'_{p,T+1}(L_s)K_{p,n,T}(I_s), \text{ and } e_{p,n,T+1}(I_s) = \left[\ I \,|\, H'_{p,n,T}(I_s)\ \right]\underline{X}_{p,T+1}(L_s) \qquad (2.6)$$

In addition, for the corresponding backward AR(p,Ms), we will have

$$R_{p,n,T}(M_s) = \sum_{i=p+n}^{T} \lambda^{T-i}\begin{bmatrix} X_{p-1,i}(M_s) \\ \hline x(i-p)\end{bmatrix}\left[X'_{p-1,i}(M_s)\right],\ D_{p,n,T}(M_s) = \sum_{i=p+n}^{T} \lambda^{T-i} X_{p-1,i}(M_s)X'_{p-1,i}(M_s),$$

$$\overline{U}_{p,T}(M_s) = \sum_{t=n+p}^{T} \lambda^{T-i}\left[\overline{\varepsilon}_p^2(t)\right],$$

and

$$R_{p,n,T}(M_s)\begin{bmatrix} G_{p,n,T}(M_s) \\ \hline 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \hline \overline{U}_{p,n,T}(M_s)\end{bmatrix},$$

where $G_{p,n,T}(M_s)$ and $X_{p-1,i}(M_s)$ are formed by removing the (p+2-m_1),...,(p+2-m_s)'th row of $G_{p,n,T}$ and $X_{p-1,i}$ respectively. We note p+2-m_j=2+i_j=o_j, thus we can easily see that $X_{p-1,i}(M_s)=X_{p-1,i}(O_s)$ and $R_{p,n,T}(M_s)= R_{p,n,T}(I_s)$.

Next, we consider a forward AR(p+1,I_s) model with a constant term of the form

$$x(t) + h_1(L_s)x(t-1) + \tau + \sum_{i=2}^{p+1} h_i(L_s)x(t-i) = \varepsilon(t), h_i(L_s) = 0, \text{as } i \in L_s,$$

where we have shifted theconstant term to the third term of the model for ease of matrix algebra operations. Suppose the model is based on the sample set {x(n), x(n+1), ..., x(T+1)}, we have

$$\underline{C}_{p+1,n,T+1}(L_s) = \sum_{i=p+1+n}^{T+1} \lambda^{T+2-i} X_{p,i-1}(O_s)X'_{p,i-1}(O_s),\ \tilde{K}_{p+1,n,T+1}(L_s) = \underline{C}^{-1}_{p+1,n,T+1}(L_s)X_{p,T+1}(O_s),$$

and $\tau_{p+1,n,T+1}(L_s) = 1 + X'_{p,T+1}(O_s)\tilde{K}_{p+1,n,T+1}(L_s).$ (2.7)

Again we consider a forward AR(p,L_s) model with constant, i.e.,

$$x(t) + h_1(L_s)x(t-1) + \tau + \sum_{i=2}^{p} h_i(L_s)x(t-i) = \varepsilon(t), h_i(L_s) = 0, \text{as } i \in L_s,$$

where we keep the constant term in the third term to assist with our algebraic manipulations. Suppose the model is fitted to the sample set {x(n+1), x(n+2),..., x(T+1)}, analogously we have

$$\tilde{K}_{p,n+1,T+1}(L_s) = \underline{C}^{-1}_{p,n+1,T+1}(L_s)X_{p-1,T+1}(O_s),$$

and $\tau_{p,n+1,T+1}(L_s) = 1 + X'_{p-1,T+1}(O_s)\tilde{K}_{p,n+1,T+1}(L_s).$ (2.8)

The matrix inversion of $R_{p,n,T}(I_s)$ provides

$$R^{-1}_{p,n,T}(I_s) = \begin{bmatrix} U^{-1}_{p,n,T}(I_s) & U^{-1}_{p,n,T}(I_s)H'_{p,n,T}(I_s) \\ H_{p,n,T}(I_s)U^{-1}_{p,n,T}(I_s) & C^{-1}_{p,n,T}(I_s) + H_{p,n,T}(I_s)U^{-1}_{p,n,T}(I_s)H'_{p,n,T}(I_s) \end{bmatrix} \tag{2.9}$$

Right multiply with $\underline{X}_{p,T+1}(L_s)$ on both sides of (2.9) and employ (2.5) and (2.6), so that we have

$$\tilde{K}_{p+1,n,T+1}(L_s) = \begin{bmatrix} 0 \\ K_{p,n,T}(I_s) \end{bmatrix} + \begin{bmatrix} I \\ H_{p,n,T}(I_s) \end{bmatrix} U^{-1}_{p,n,T}(I_s)e_{p,n,T+1}(I_s). \tag{2.10}$$

Again we right multiply the transpose of (2.10) with $\underline{X}_{p,T+1}(L_s)$, employ (2.5), and add 1 to both sides, so that we now establish

$$\tau_{p+1,n,T+1}(L_s) = \tau_{p,n,T}(I_s) + e_{p,n,T+1}(I_s)U^{-1}_{p,n,T}(I_s)e_{p,n,T+1}(I_s). \tag{2.11}$$

Analogously we can have

$$\tilde{K}_{p+1,n,T+1}(L_s) = \begin{bmatrix} \tilde{K}_{p,n+1,T+1}(L_s) \\ 0 \end{bmatrix} + \begin{bmatrix} G_{p,n,T}(M_s) \\ I \end{bmatrix} \overline{U}^{-1}_{p,n,T}(M_s)\bar{e}_{p,n,T+1}(M_s),$$

$$\tau_{p+1,n,T+1}(L_s) = \tau_{p,n+1,T+1}(L_s) + \bar{e}_{p,n,T+1}(M_s)\overline{U}^{-1}_{p,n,T}(M_s)\bar{e}_{p,n,T+1}(M_s),$$

where $\bar{e}_{p,n,T+1}(M_s) = \left[G'_{p,n,T}(M_s) \,|\, I \right] X_{p,T+1}(M_s)$.

From Penm *et al* (1995), if we permute the first row and the second row of the $\tilde{K}_{p+1,n,T+1}(I_s)$, the resulting vector is the $K_{p+1,n,T+1}(I_s)$ associated with a forward AR(p+1,Ls) model of the form

$$x(t) + \delta + h_1(L_s)x(t-1) + \sum_{i=2}^{p+1} h_i(L_s)x(t-i) = \varepsilon(t), \quad h(i, L_s) = 0, \text{as } i \in L_s,$$

which means that $K_{p+1,n,T+1}(I_s) = P_{p+1}\tilde{K}_{p+1,n,T+1}(I_s)$, where P_{p+1} is a permutation matrix of the form:

$$\begin{bmatrix} 0 & 1 & 0 & \vdots & 0 \\ 1 & 0 & 0 & \vdots & 0 \\ 0 & 0 & 1 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \vdots & 1 \end{bmatrix}.$$

Note that if there is a consecutive set of k deleted lags beginning at lag 1 in the forward AR (p, I_s) model fitted using the sample {x(1), ..., x(T+1)}, we have

$$K_{p,nT+1}(I_s) = K_{p-k,n,T+1-k}(I_k) \text{ and } \tau_{p,n,T+1}(I_s) = \tau_{p-k,n,T+1-k}(I_k), \tag{2.12}$$

where I_s contains i_1, ..., i_k, ..., i_s, and i_j=j, j=1,2, ..., k, and I_k contains i_{k+1}, ..., i_{s-k}.

In implementing the recursions using the exact-windowed case, the starting point of the data range has been shifted from n to n+1 in the following steps:

$$K_{p,n,T}(I_s) \to K_{p+1,n,T+1}(I_s) \to K_{p,n+1,T+1}(I_s).$$

If n=1, this means that the starting point of the data range has been shifted from point 1 to point 2. That is $K_{p,1,T}$ (I_s) to $K_{p,2,T+1}$ (I_s).

However in the forgetting factor exclusive algorithm proposed in Penm *et al* (1995), the starting point of the data range has always been fixed to point 1. Thus the following steps apply:

$K_{p,1,T}(I_s)$ to $K_{p+1,1,T+1}(I_s)$ to $K_{p,1,T+1}(I_s)$ [not $K_{p,2,T+1}(I_s)$].

Therefore the current forgetting-factor inclusive algorithm is different from the algorithm without the forgetting factor proposed in Penm *et al* (1995). The current algorithm, with a constant term or without a constant term, will shift the beginning point to compute the Kalman gain vector, K, from n to n+1 (not fixed), but the algorithm proposed in Penm *et al* (1995) use a fixed starting point which is always Point 1.

This change from the beginning point n to n+1 of the data range also applies to the angular variable, τ. The current algorithm, with a constant term or without a constant term, will shift the beginning point to compute the angular variable from n to n+1 (not fixed), but the algorithm proposed in Penm *et al* (1995) uses a fixed starting point which is always Point 1.

Now we turn to the SDL modelling, and the following relations at t=T+1 have been established.

$$R_{p-1,n,T+1}(I_s)A_{p,n,T+1}(I_s) = -r_{p-1,n,T+1}(I_s), \quad (2.13)$$

where $R_{p-1,n,T+1}(I_s) = \lambda R_{p-1,n,T}(I_s) + X_{p-1,T+1}(I_s)X'_{p-1,T+1}(I_s)$,

and $r_{p-1,n,T+1}(I_s) = \lambda r_{p-1,n,T}(I_s) + X_{p-1,T+1}(I_s)y(T+1)$

We compare (2.13) to the relations at t=T, the following time update recursions for SDL modelling can be developed.

$$\theta_{p,n,T+1}(I_s) = y(T+1) + A'_{p,n,T}(I_s)X_{p-1,T+1}(I_s) \quad (2.14)$$

$$\eta_{p,n,T+1}(I_s) = \theta_{p,n,T+1}(I_s)\tau^{-1}_{p,n,T+1}(I_s) \quad (2.15)$$

$$A_{p,n,T+1}(I_s) = A_{p,n,T}(I_s) - K_{p,n,T+1}(I_s)\ \eta_{p,n,T+1}(I_s) \quad (2.16)$$

$$\Omega_{p,T+1}(I_s) = \lambda\Omega_{p,T}(I_s) + \eta_{p,n,T+1}(I_s)\ \theta_{p,n,T+1}(I_s) \quad (2.17)$$

In addition, the forward-time update algorithm from T to T+1 is summarised in Table 1.

3. An illustration

Assuming two sets of data, $\{x_1,x_2,\ldots,x_{97},\ldots\}$ and $\{y_1,y_2,\ldots,y_{97},\ldots\}$, suppose we wish to employ the above proposed forward-time update recursions for all possible subset AR models and SDL models from (n,T)=(1,97) to (n,T)=(1,98), where it is assumed that the maximum lag for AR models and SDL models is P=4 for illustration purposes. Also, $H_{p,n,T}(I_s)$, $U_{p,n,T}(I\ )$, $K_{p,n,T}(I_s)$, $\tau_{p,n,T}(I_s)$, $G_{p,n,T}(M_s)$, $X_{p,T+1}(O_s)$, . $\overline{U}_{p,n,T}(M_s)$, $A_{p,n,T}(I_s)$, $X_{p-1,T+1}(I_s)$ and $\Omega_{p,T}(I_s)$ are available at time T = 97, By employing the equations (2.18a)-(2.18m), we can update the matrices $H_{p,n,T}(I_s)$ and $G_{p,n,T}(M_s)$ and parameters $U_{p,n,T}(I_s)$ and $\overline{U}_{p,n,T}(M_s)$ for each reciprocal integer pair of the AR(p, I_s) and AR(p, M_s) at T=98. Both $K_{p,n,T}(I_s)$ and $\tau_{p,n,T}(I_s)$ at T=98 corresponding to the forward autoregressions that include lag 1 can also be acquired from the recursions (2.18e) and (2.18h). Figure 2 illustrates the situation, single lines show the recursions. Further, by employing the equations (2.18n)-(2.18q), we can update the matrix, $A_{p,n,T}(I_s)$, and parameter $\Omega_{p,T}(I_s)$ for each SDL model at T=98.

Both $K_{p,n,T}(I_s)$ and $\tau_{p,n,T}(I_s)$ at T=98 corresponding to the forward autoregressions that exclude lag 1 may be obtained from (2.12), i.e. the quantities $K_{p,n,T}(I_s)$ and $\tau_{p,n,T}(I_s)$ will be identical at T<98 for the forward autoregressions specified by double lines in Figure 2. SDL models which parameters will be updated by using both K and τ acquired are also shown in Figure 2. At this point, we can carry out the forward time-update recursions from T=98 to 99, T=99 to 100, and so on.

To determine the optimal SDL model at each time instant, we utilise the order selection criterion suggested by Hannan and Deistler (1988). From now on, we will use MHQC as an abbreviation for this criterion, which is defined by

MHQC= $\log(\text{estimated residual variance}) + [2 \log\log f(T)/f(T)]N$,

where $f(T)= \sum_{t=p-1+n}^{T} \lambda^{T-t}$ is the effective sample size (see Hannan and Deistler, 1988), and N the number of functionally independent parameters. The optimal model selected is the one with the minimum value of MHQC.

4. Features of the current forgetting factor inclusive algorithm using the exact-windowed case

As described in Section 3, the current forgetting factor inclusive algorithm using the exact-windowed case is quite different from the algorithm using the pre-windowed case (Penm *et al*, 1995). A significant difference is as follows:

The current algorithm uses the exact-windowed case to estimate the parameters of a DL model. For simplicity we consider the following DL(p) model without forgetting factor:

$$y(t) + \sum_{i=0}^{p} a_i x(t-i) = \varepsilon(t) \qquad (3.1)$$

For two given sets of data $\{x(n), x(n+1), \ldots, x(T)\}$ and $\{y(n), y(n+1), \ldots, y(T)\}$, the model (3.1) using the exact-windowed model becomes

t	model
T	$y(T) + a_0 x(T) + \cdots + a_p x(T - p) = \varepsilon(T)$
T-1	$y(T - 1) + a_0 x(T - 1) + \cdots + a_p x(T - p - 1) = \varepsilon(T - 1)$
.	…
.	…
.	…
n+p	$y(n + p) + a_0 x(n + p) + \cdots + a_p x(n) = \varepsilon(n + p)$ (3.2)

Therefore, all observations used in (3.2) are available in the observed sample {x(n),…,x(T)}. However the algorithm proposed in Penm *et al* (1995) use the pre-windowed case, which will use unobserved observations to undertake estimation. For a given observed sample {x(1),…,x(T)}, the model using the pre-windowed case becomes

t	model
T	$y(T) + a_0 x(T) + \cdots + a_p x(T - p) = \varepsilon(T)$
T-1	$y(T - 1) + a_0 x(T - 1) + \cdots + a_p x(T - p - 1) = \varepsilon(T - 1)$
.	…
.	…
.	…
p	$y(p) + a_0 x(p) + \cdots + a_p x(0) = \varepsilon(p)$ (3.3)
p-1	$y(p - 1) + a_0 x(p - 1) + \cdots + a_p x(-1) = \varepsilon(p - 1)$
…	
1	$y(1) + a_0 x(1) + \cdots + a_p x(-p + 1) = \varepsilon(1)$.

The pre-windowed case therefore needs observations prior to time 1, {x(0),x(-1),..x(-p+1)}, and so consider x(0), x(-1),…x(-p+1) all equal to zero. Since the pre-windowed case has x(-p+1) as the starting point of the data range, and consider each earlier unseen observation is equal to zero, this case is definitely different from the exact-windowed case. The implication is that significantly different parameter estimates can be obtained, in particular, in small sample cases.

Also, the forgetting factor has been incorporated into K and τ in the current forgetting-factor inclusive algorithm. Therefore the current forgetting-factor inclusive algorithm is different from the algorithm without the forgetting factor proposed in Penm *et al* (1995), and thus will produce different estimation results, in particular in small sample cases.

Further, the current algorithm is a coefficient-based time update algorithm, which can detect evolutionary changes in model structures. However the proposed lattice order update algorithms in Haykin (1996) are residual-based algorithms, which undertake recursions, moving from low-order models to high-order models, so no evolutionary changes are captured through parameter updating. Therefore the focus of Haykin (1996) is not on the evolution over time of the parametric structure of the system. The focus of Haykin (1996) is only on checking at each time point of the data how the order and the subset make-up of the DL changes, i.e. on lag length and subset lag inclusion or exclusion.

5. The causal relationship between futures and the underlying assets

Two illustrations are presented to demonstrate the practical use of the algorithm. In the first illustration the relationship between Australian's ten-year Commonwealth Treasury bond futures contracts (TBFC) and the underlying bond assets is examined.

In the second one, the relationship between Australian's All Ordinaries Share Price Index (AOI) and Australian's All Ordinaries Share Price Index futures (SPI) is investigated. A futures contract is one of the most important hedging instruments for the underlying asset. Both bond and stock index futures have many attractive hedging benefits for a trader who wishes to trade the underlying asset portfolio corresponding to the index. In Australia, Commonwealth Government bonds are the most important fixed interest securities, in line with government borrowing. The main Commonwealth Government bond security is the 10-year Commonwealth Treasury bond (TB), which is traded on the Australian Stock Exchange. The Sydney Futures Exchange offers a futures contract on the ten-year Treasury bond. This contract is available on a quarterly expiry date and is known as the ten-year Commonwealth TBFC.

Further, the main stock market indicator is the AOI. The index is calculated on the basis of market capitalisation of the constituent stocks traded on the Australian Stock Exchange. The Sydney Futures Exchange offers a futures contract on the AOI. This contract is available on a quarterly expiry date and is known as the SPI Futures Contract.

There is already a considerable literature examining the relationship between futures and the underlying asset market prices. The literature has examined either theoretical relationships between the markets through models such as the cost-of-carry (see Brailsford and Hodgson, 1997), or the causality between the markets through lead-lag relationships, cointegration tests or bivariate spillover models (see Chan, 1992; Martens *et al*, 1996). The general findings confirm a strong causality between the markets (see Wahab and Lashgari, 1993; Abhyankar, 1995). This relationship is not unexpected given the pricing relationship between the markets and the fact that the basis reduces to zero at the maturity of the futures contract. However there has been debate about the direction of causality, with the evidence generally indicating that the futures market leads the stock market. In particular, Chan (1992) has examined the lead-lag relation between returns of the Major Market cash index and returns of the Major Market Index futures and S&P 500 futures. His results indicate that the futures price is a leading indicator for the spot, when stock prices move together under market-wide movements. Tse (1995) has studied the causal relation between stock index futures and cash index prices in Japan, and documents that futures prices cause cash index prices.

For the first illustration, data on the TB, TBFC, AOI and SPI are sampled daily between 18 March 2003 and 10 September 2003. Both the TB and AOI data are observed as the daily market closing value whereas both the TBFC and SPI data are observed as the last traded price on each day in the September 2003 contract. Graphs of log TB, log TBFC, log AOI and log SPI in first differences are shown in Figures 3, 4, 5 and 6 respectively. To test for the unit-roots for each plotted series, Microfit 4.0 is used to carry out the augmented Dickey-Fuller (ADF) unit root test. The 95 per cent critical values for each test computed using the response surface estimates, indicate that whereas all log TB, log TBFC, log AOI and log SPI are non-stationary, their data in first differences, $\Delta\log\text{TB}$, $\Delta\log\text{TBFC}$, $\Delta\log\text{AOI}$ and $\Delta\log\text{SPI}$, are stationary for the period 18 March 2003 to 11 August 2003 (T= 105). Furthermore, $\Delta\log\text{TB}$, $\Delta\log\text{TBFC}$, $\Delta\log\text{AOI}$ and $\Delta\log\text{SPI}$ continue to be stationary for the extended period from 18 March 2003 to 11 August 2003.

The algorithm developed in Section 2 is used to model the relationships between the TB and TBFC. In detecting the causal relationship from log TBFC to log TB, the variables used are y(t) = log TBFC and x(t) = log TB. As discussed above, neither log TBFC nor log TB are stationary. Therefore, the forgetting factor, λ, is incorporated to allow for the presence of non-stationarity. To begin it is assumed P=16, which corresponds to a three-week period (i.e. 15 business days). The evolutionary kernel-based SDL recursions described above are then used to select the 'optimal' specification of the discrete lag models at T=105, 106,…, 112.

For the causal relationship from log TB to log TBFC, the optimal discrete lag models with λ=0.999 and 0.985 are presented in Table 2. To check the adequacy of each optimal model fit, the strategy suggested in Tiao and Tsay (1989) and Brailsford *et al* (2001) is used, with the proposed Penm and Terrell (1984) algorithm applied to test each residual series.[1] The results in table 1 support the hypothesis that each residual series is a white noise process. These optimal models are then used as the benchmark models for analysing the causal relationships. The lower value of λ is consistent with strong persistence in market price fluctuations. For brevity, only the results obtained by the MHQC are presented. For cases where λ is less than 0.985, the selected models are not reported due to the small effective sample size (<50).

To assess the causality from log TB to log TBFC, a SDL model for λ =0.999 and 0.985 with lags (0, 2, 7) was selected by the MHQC at T=105, 106,…, 110. At T=111, the lag structure selected changes to (0, 2). In this case the predicted output, log TBFC, is related to the current and previous inputs of log TB. Also, the lag 0 indicates instantaneous causality from TBFC and TB. These results indicate that instantaneous and direct causal relationships exist from the bond market to the futures market, even when emphasis is placed on recent data.

---

[1] In subset time-series modeling Tiao and Tsay (1989) propose an algorithm using the crit(m,j) criterion to select the subset autoregressive moving average process. After the final model is selected, their algorithm is then applied to the residual series to test whether this series is a white noise process.

For the causal relationship from log TBFC to log TB, Table 3 shows the optimal discrete lag models with λ=0.999 and 0.985. These results strongly support the existence of instantaneous and bi-directional causal relationships between the Australian bond and index futures markets. Also, the results in table 3 support the hypothesis that each residual series is a white noise process

For the second illustration, Table 4 presents a subset model for λ =0.999 with lags (0, 1, 5), to assess the causality from log SPI to log AOI, which was selected by the MHQC at T=105,…, 109. At T=110, the lag structure selected changes to (0, 5). In this case, the model specifications indicate that log AOI is related to the current and previous inputs of log SPI. Also, the lag 0 indicates instantaneous causality between AOI and SPI. For λ =0.985, The optimal model with lags (0, 5) is selected at T=105,…, 112. These results indicate that both instantaneous and direct causal relationships exist from the futures market to the stock market when emphasis is placed on recent data.

For the causal relationship from log AOI to log SPI, Table 5 shows the optimal discrete lag models with λ=0.999 and 0.985. These results strongly support the existence of instantaneous causal and bidirectional feedback relationships between the Australian stock and index futures markets. These conclusions are generally consistent with those reported elsewhere in similar markets by Wahab and Lashgari (1993), and Abhyankar (1995).

In general these outcomes can be explained by reference to transaction costs, time delays in computing the index, execution costs, and measurement errors (see Chan, 1992). In addition to speculators, some other investors, particularly institutional investors, participate in the futures market for hedging purposes. Usually they take opposite positions in the underlying asset market and the futures market at the same time, in order to hedge their exposure. Since they participate in both markets, price information will flow between the two markets. Therefore the finding of bidirectional causal relationships between these two markets is consistent with our prior hypotheses.

Note there are other financial variables which could play a significant role in the above asset market analysis. This application merely demonstrates the usefulness of the proposed evolutionary kernel-based SDL recursive algorithm in time-series analysis.

6. Summary

In this paper an evolutionary kernel-based time update algorithm has been presented to recursively estimate SDL models with a forgetting factor, using the exact-windowed case. Two financial market illustrations to demonstrate the use of the proposed evolutionary algorithm are provided. Compared to the pre-windowed case, the exact-windowed case utilises only the available observations, without any assumption on unseen observations, to estimate model coefficients. We prefer evolutionary parameter updating algorithms because they allow users to update subset time-series models at consecutive time instants, and can show evolutionary changes detected in model structures. This is in contrast to residual-based order update algorithms (see Haykin, 1996), which undertake recursions, moving from low-order models to high-order models, whereby no evolutionary changes are captured through parameter updating.

References

Abhyankar, A. (1995). Return and volatility dynamics in the FTSE-100 stock index and stock index futures markets, *Journal of Futures Markets,* 15(4), 457–488.

Brailsford, T.J. and Hodgson, A. (1997). Mispricing in stock index futures: A re-examination using the SPI, *Australian Journal of Managemen*t, 22, 21-43.

Brailsford, T.J., Penm, J. and Terrell, R.D. (2001). The adjustment of the Yule-Walker relations in VAR Modeling: The impact of the Euro on the Hong Kong stock market, *Multinational Finance Journal*, 5(1),35-58.

Brailsford, T.J., Penm, J., Terrell, R.D. and Wu, S. (2005). A new approach of kernel bandwidth applications for time-series using the example of the prediction of the euro's exchange rate, Accepted in 2004, *Research in Finance* 22, 2005.

Chan, K. (1992). A further analysis of the lead-lag relationship between the cash and stock index futures market, *Review of Financial Studies*, 5, 123-151.

Gijbels, I., Pope, A. and Wand, M. P. (1999). Understanding exponential smoothing via kernel regression, *J. R. Stat. Soc. Ser. B* 61, 39–50, 1999

Hannan, E.J. and Deistler, M. (1988). *The statistical theory of linear systems*, John Wiley and Sons, New York.

Haykin, S. (1996). *Adaptive filter theory*. Upper Saddle River, N.J., Prentice Hall.

Holmes, J.M. and Hutton, P.A. (1989). 'Optimal' model selection when the true relationship is weak and occurs with a delay, *Economics Letters,* 30, 333-339.

Kalouptsidis, N., Carayannis, G. and Manolakis, D. (1984). A fast covariance type algorithm for sequential least-squares filtering and prediction, *IEEE Transactions on Automatic Control*, 29, 8, 752-755.

Martens, M., Kofman, P. and Vorst, T. (1996). A threshold error correction model for intraday futures and index returns, Working paper, Erasmus University Rotterdam.

Mittnik, S. (1989). Forecasting with balanced state space representations of multivariate discrete lag models, *Journal of Forecasting*, 9, 207-218.

Penm, J.H.W., and Terrell, R.D. (1984). Multivariate subset autoregressive modeling with zero constraints for detecting causality. *Journal of Econometrics* 3: 311-30.

Penm, J.H., Penm, J.H.W. and Terrell, R.D. (1995). A note on the sequential fitting of subset autoregressions using the prewindowed case, *IEEE Transactions on Signal Processing*, 43, 322-27.

Refenes, A., Abu-Mustafa, Y., Moody, J. and Weigend, A. (1996). Eds., *Neural Networks in Financial Engineering, vol. 2 of Progress in Neural Processing,* World Scientific Publishing, Co., London.

Tiao, G.C., and Tsay, R.S. (1989). Model specification in multivariate time-series. *Journal of the Royal Statistical Society B* 51: 157-213.

Tse Y.K. (1995). Lead-lag relationship between spot index and futures prices of the Nikkei stock average, *Journal of Forecasting*, 14, 553-563.

Wahab, M. and Lashgari, M. (1993). Price dynamics and error correction in stock index and stock index futures markets: a cointegration approach. *Journal of Futures Markets*, 13, 711–742.

Figure 1: A four-variable tree diagram which illustrates the reciprocal integer pairs of subset AR models up to and including lag length P=4. Of note, numerals denote particular lags in a forward AR and numerals in italics denote such leads in a backward AR.

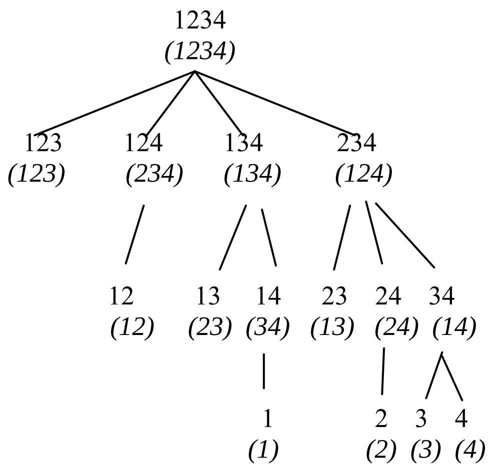

Figure 2: Both K and τ acquired by using the time updating approach for an AR(4) variable case and the associated SDL variable case.

a) Using the recursions (2.18e) and (2.18h)

T=97	T=98
123 →	1234 (0123)
12 →	123 (012)
13 →	124 (013)
23 →	134 (023)
1 →	12 (01)
2 →	13 (02)
3 →	14 (03)
0 →	1 (0)

b) Using the relations $K_{p,n,T+1}(I_s) = K_{p-k,n,T+1-k}(I_k)$ and $\tau_{p,n,T+1}(I_s) = \tau_{p-k,n,T+1-k}(I_k)$

T=95	T=96	T=97	T=98
		123 ⇒	234 (123)
		12 ⇒	23 (12)
		13 ⇒	24 (13)
	12 ⇒		34 (23)
		1 ⇒	2 (1)
	1 ⇒		3 (2)
1 ⇒			4 (3)

. Only the forward AR of each pair has been listed.

. Underline numerals denote particular lags in a SDL whose parameters will be updated by using both K and τ acquired

Figure 3

Log TB in first differences, daily: 18 March 2003 to 10 September 2003

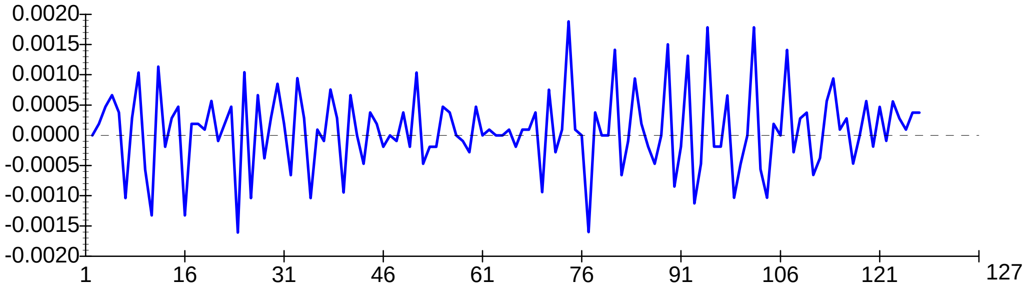

Figure 4

Log TBFC in first differences, daily: 18 March 2003 to 10 September 2003

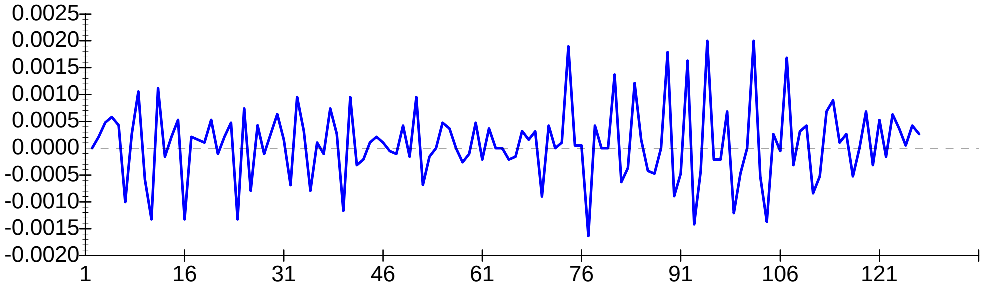

Figure 5

Log AOI in first differences, daily: 18 March 2003 to 10 September 2003

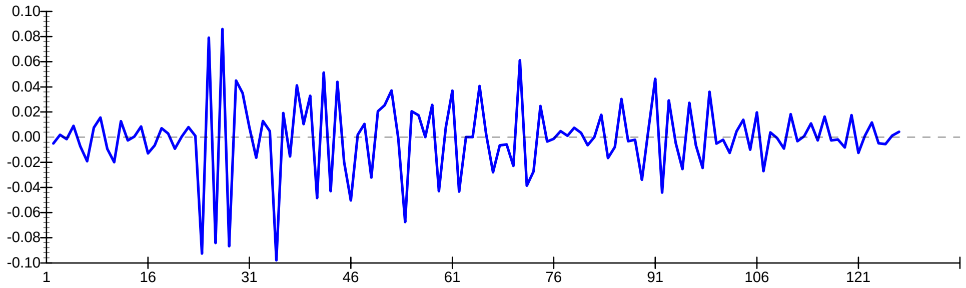

Figure 6

Log SPI in first differences, daily: 18 March 2003 to 10 September 2003

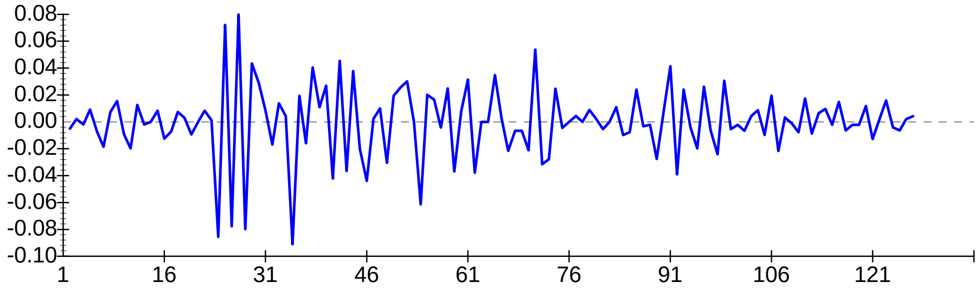

Table 1: The forward-time update recursions from T to T+1 for subset AR forgetting-factor inclusive models with intercept variable

$H_{p,n,T}(I_s)$, $U_{p,n,T}(I\ )$, $K_{p,n,T}(I_s)$, $\tau_{p,n,T}(I_s)$, $G_{p,n,T}(M_s), \lambda$,
$X_{p,T+1}(O_s)$, $X_{p-1,T+1}(I_s)$, $\bar{U}_{p,n,T}(M_s)$, $A_{p,n,T}(I_s)$, $\Omega_{p,T}(I_s)$, and y(T+1) are available

Recursions:

$$e_{p,n,T+1}(I_s) = [1 \vdots H'_{p,n,T}(I_s)]\, X_{p,T+1}(O_s) \quad (2.18a)$$

$$\varepsilon_{p,n,T+1}(I_s) = e_{p,n,T+1}(I_s)\tau_{p,n,T}{}^{-1}(I_s) \quad (2.18b)$$

$$H_{p,n,T+1}(I_s) = H_{p,n,T}(I_s) - K_{p,n,T}(I_s)\ \varepsilon_{p,n,T+1}(I_s) \quad (2.18c)$$

$$U_{p,n,T+1}(I_s) = \lambda U_{p,n,T}(I_s) + \varepsilon_{p,n,T+1}(I_s)\ e_{p,n,T+1}(I_s) \quad (2.18d)$$

$$\tilde{K}_{p+1,n,T+1}(L_s) = \begin{bmatrix} 0 \\ K_{p,n,T}(I_s) \end{bmatrix} + \begin{bmatrix} 1 \\ H_{p,n,T}(I_s) \end{bmatrix} U_{p,n,T}{}^{-1}(I_s)\ e_{p,n,T+1}(I_s) \quad (2.18e)$$

partition $\tilde{K}_{p+1,n,T+1}(L_s) = \begin{bmatrix} D \\ \hline d \end{bmatrix}$

$$\bar{e}_{p,n,T+1}(M_s) = \lambda \bar{U}_{p,n,T}(M_s)d \quad (2.18f)$$

$$\tilde{K}_{p,n+1,T+1}(L_s) = D - G_{p,n,T}(M_s)d \quad (2.18g)$$

$$\tau_{p+1,n,T+1}(L_s) = \tau_{p,n,T}(I_s) + e_{p,n,T+1}(I_s)U_{p,n,T}{}^{-1}(I_s)\ e_{p,n,T+1}(I_s) \quad (2.18h)$$

$$\tau_{p,n+1,T+1}(L_s) = \tau_{p+1,n,T+1}(L_s) - \bar{e}_{p,n,T+1}(M_s)d \quad (2.18i)$$

$$\bar{\varepsilon}_{p,n,T+1}(M_s) = \bar{e}_{p,n,T+1}(M_s)[\tau_{p,n+1,T+1}(L_s)]^{-1} \quad (2.18j)$$

$$\bar{U}_{p,n,T+1}(M_s) = \lambda \bar{U}_{p,n,T}(M_s) + \bar{\varepsilon}_{p,n,T+1}(M_s)\bar{e}_{p,n,T+1}(M_s) \quad (2.18k)$$

$$G_{p,n,T+1}(M_s) = G_{p,n,T}(M_s) - \tilde{K}_{p,n+1,T+1}(L_s)\bar{\varepsilon}_{p,n,T+1}(M_s) \quad (2.18l)$$

$$K_{p+1,n,T+1}(I_s) = P_{p+1}\tilde{K}_{p+1,n,T+1}(I_s) \quad (2.18m)$$

$$\theta_{p,n,T+1}(I_s) = y(T+1) + A'_{p,n,T}(I_s)X_{p-1,T+1}(I_s) \quad (2.18n)$$

$$\eta_{p,n,T+1}(I_s) = \theta_{p,n,T+1}(I_s)\tau^{-1}_{p,n,T+1}(I_s) \quad (2.18o)$$

$$A_{p,n,T+1}(I_s) = A_{p,n,T}(I_s) - K_{p,n,T+1}(I_s)\ \ \eta_{p,n,T+1}(I_s) \quad (2.18p)$$

$$\Omega_{p,T+1}(I_s) = \lambda\Omega_{p,T}(I_s) + \eta_{p,n,T+1}(I_s)\ \ \theta_{p,n,T+1}(I_s) \quad (2.18q)$$

Table 2

The SDLs selected by MHQC for detecting the causal relationship from TB to TBFC

Sample size (T)	Non-zero lag structure for y(t)=log TBFC and x(t)=log TB		Pattern of causality[a]
	λ=0.999	λ=0.985	
105,106,107,108,109, 110	0 2 7	0 2 7	log TB → log TBFC
111,112	0 2	0 2	log TB → log TBFC

(a) w → z: w causes z directly and instantaneously.

(b) To check the adequacy of the model fit, the results in table 2 support the hypothesis that the residual series of each selected model is a white noise process.

Table 3

The SDLs selected by MHQC for detecting the causal relationship from TBFC to TB

Sample size (T)	Non-zero Lag Structure for y(t)=log TB and x(t)=log TBFC		Pattern of causality[a]
	λ=0.999	λ=0.985	
105,106,107,108,109, 110,111,112	0 2	0 2	Log TBFC → log TBI

(a) w → z: w causes z directly and instantaneously.

(b) To check the adequacy of the model fit, the results in table 2 support the hypothesis that the residual series of each selected model is a white noise process.

Table 4

The SDLs selected by MHQC for detecting the causal relationship from SPI to AOI

Sample size(T)	Non-zero lag structure for y(t)=log AOI and x(t)=log SPI		Pattern of causality[a]
	λ=0.999	λ=0.985	
105,106,107,108,109	0 1 5	0 5	log SPI → log AOI
110,111,112	0 5	0 5	log SPI → log AOI

(a) w → z: w causes z directly and instantaneously.

(b) To check the adequacy of the model fit, the results in table 2 support the hypothesis that the residual series of each selected model is a white noise process.

Table 5

The SDLs selected by MHQC for detecting the causal relationship from AOI to SPI

Sample size (T)	Non-zero Lag Structure for y(t)=log SPI and x(t)=log AOI		Pattern of causality[a]
	λ=0.999	λ=0.985	
105,106,107,108,109	0 1 5	0 1 5	log AOI → log SPI
110,111,112	0 5	0 5	log AOI → log SPI

(a) w → z: w causes z directly and instantaneously.

(b) To check the adequacy of the model fit, the results in table 2 support the hypothesis that the residual series of each selected model is a white noise process.