The Adjustment of the Yule-Walker Relations in VAR Modeling: The Impact of the Euro on the Hong Kong Stock Market*

Timothy J. Brailsford
The Australian National University, Australia

Jack H.W. Penm
The Australian National University, Australia

R. Deane Terrell
The Australian National University, Australia

Vector autoregressive models are increasingly being used in the analysis of relationships within and between financial markets. In such models, there are circumstances that require zero entries in the coefficient matrices. Such circumstances can be particularly relevant in the context of markets with special characteristics, such as emerging economies. This paper shows that a direct extension of the use of the Yule-Walker relations for fitting vector autoregressive models with zero-non-zero patterned coefficient matrices is inconsistent with statistical procedures as the resultant estimated variance-covariance matrix of the white noise disturbance process becomes non-symmetric. This inconsistency can cause a breakdown when testing financial theory. The paper provides a consistent adjustment which fits with the theory. The practical use of the adjustment is demonstrated in a vector system comprising variables from the Hong Kong stock market and foreign exchange markets (JEL C13, C32, C63, G10, G15).

Keywords: foreign exchange market, time series, VAR models, Yule-Walker relations.

I. Introduction

Vector autoregressive (VAR) models represent an advance in the analysis of time series. These models provide a device that has proved to be a more computationally efficient tool, and therefore less costly,

*The authors acknowledge the financial assistance provided by the Australian Research Council Small Grants Scheme, and the comments of two anonymous referees and the editor.

(*Multinational Finance Journal*, 2001, vol. 5, no. 1, pp. 35–58)

DOI: 10.17578/5-1-2

than conventional financial and econometric time series techniques. In recent years the use of VAR models as a means of modeling financial time series has become common. In particular, VAR modeling has been increasingly employed to examine relationships in stock markets. For instance Eun and Shim (1989) estimate a VAR using index returns on nine stock markets to examine interactions among the markets. In the context of emerging stock markets, Bekaert et al. (1999) estimate a VAR using capital flows, equity returns, dividend yields and interest rates to examine the extent to which lower interest rates contribute to increased capital flows. In a similar study, Froot et al. (1998) employ VAR estimation to examine the relationship between capital flows and equity returns in emerging markets.

The use of VAR models for econometric research has in part been driven by the desire to provide users with a relatively simple forecasting procedure accessible to non-specialists. However early researchers realised that heavy parameterisation of their VAR models resulted in poor *ex-ante* forecasting performance. Their proposed procedures rested on the assumption that the coefficient matrices of the VAR model had all non-zero entries. In effect, the assumption of non-zero entries restricts the range of possible model specifications. Further, if the true underlying VAR process has zero entries in its structure, then sub-optimal model design induced by assuming a full-order structure can produce misleading inferences and inferior projections. Consequently, models have been developed that allow for zero entries in the coefficient matrices such as a zero-non-zero (ZNZ) patterned structure. However, implementation of a ZNZ structure in a VAR is difficult given the large number of parameters and possibilities. That is, in the absence of an effective approach to find the optimal model, relaxation of the assumption of non-zero entries is problematic.

The issue is also relevant when investigating causality.[1] Optimal VAR models with ZNZ patterned coefficient matrices can also be used as a basis for detecting Granger-causality and the instantaneous

1. Examples of financial and economic variables that have been tested for Granger-causal relations include volatility transmission (Bhattacharya et al. 2000), stock markets and foreign exchange (Bekaert and Hodrick 1992), volatility of stock market returns (Whitelaw 1994), monetary policy and the stock market (Thorbecke 1997), cross-market relationships (Malliaris and Urrutia 1992) and interest rates (Hassapis et al. 1999).

causality among time series variables.[2] Granger-causality and the instantaneous causality have been defined by Granger (1969), and are based entirely on the predictability of the objective variables such that the definitions make no explicit use of economic and financial laws to provide *a priori* restrictions on the structure. As such, the imposition of non-zero restrictions is inconsistent with their basic theory. Further, the growing reliance on VAR models and their use in testing for causality is limited to the extent that non-zero entries are initially assumed.

One approach to select the optimal ZNZ patterned VAR model has been advocated by Penm and Terrell (1984), and it centres on their development of a search algorithm using the Yule-Walker relations in conjunction with model selection criteria. However, that approach does not examine the estimation of the residual variance-covariance relation, rather only the Yule-Walker coefficient relations are considered. In this paper, an approach is provided that considers the variance-covariance relation within the Yule-Walker relationship and leads to an effective approach to identify the optimal ZNZ model within the context of a VAR system. The development of the approach in this paper is a significant contribution given the extent to which prior literature has imposed non-zero restrictions. The paper provides an empirical application of the approach that focuses on the Hong Kong stock market. The application highlights the usefulness and practicality of the approach.

The remainder of this paper is organised as follows. Section II provides the detailed background and describes the use of the Yule-Walker relations for fitting VAR models. Section III shows the problems and theoretical inconsistency that arise by using a two-variable VAR example. Section IV presents the new approach and contains the main contribution of the paper. In section V, an application concerning the Hong Kong stock market and the foreign exchange market is presented and concluding remarks are provided in section VI.

II. Using the Yule-Walker Relations for Fitting VAR Models

In this section, the fitting of a VAR with zero coefficient restrictions is

2. Of note, recent cointegration work suggests that, if cointegrating relations exist between the variables, then the use of the vector error-correction model, which is associated with the VAR model with unit roots, may be more effective for testing Granger-causality.

presented. First, let $u(t)=\{u_1(t),u_2(t),\ldots,u_m(t)\}'$ be a zero mean, wide-sense stationary time series of dimension m. Consider the vector *AR (p)* model of the form:

$$\sum_{k=0}^{p} A_k u(t-k)=\varepsilon(t), \tag{1}$$

where $A_0=I$, A_k, k=1,..., p are the mxm parameter matrices and $\varepsilon(t)$ is an mx1 stationary vector process with E{$\varepsilon(t)$}=0, and thus:

$$E\{\varepsilon(t)\varepsilon'(t-k)\}\begin{cases}=V \text{ as } k=0\\ =0 \text{ as } k>0\end{cases}. \tag{2}$$

The sample lag covariance matrices,

$$\Gamma_k=\frac{1}{N}\sum_{t=1}^{N-k} u(t+k)u'(t), \tag{3}$$

obey the following Yule-Walker relations. The Yule-Walker coefficient relations are:

$$\Gamma_j+\sum_{k=1}^{p}\hat{A}_k\Gamma_{j-k}=0 \quad (j=1,\ldots,p). \tag{4}$$

The Yule-Walker residual variance-covariance relation is:

$$\Gamma_0+\sum_{k=1}^{p}\hat{A}_k\Gamma_{-k}=\hat{V}, \tag{5}$$

where $\Gamma_k=\Gamma'_{-k}$; N is the sample size, $\hat{A}_k$ and $\hat{V}$ are the estimates of A_k and V respectively, and $|\hat{V}|$ is described as the generalised residual sum of squares.

In a full-order VAR model, all possible models with zero coefficient

elements are neglected. For example in a bivariate VAR model when $p = 5$, the coefficients, A_1, A_2, up to and including A_5 are assumed non-zero. However there are $2^{20} = 65{,}536$ possible models in this example. Thus a large number of possible models will be ignored under the restriction of non-zero coefficients. More important, if the true underlying VAR process has a zero-non-zero (ZNZ) patterned structure, a sub-optimal model design such as a full-order structure can produce less powerful and therefore potentially misleading inferences and inferior projections.

Penm and Terrell (1984) have proposed a search algorithm, using the Yule-Walker relations for fitting VAR models in conjunction with model selection criteria, to select the optimal ZNZ patterned VAR models. Background information on the fitting of VAR models using the Yule-Walker relations is presented in appendix 1. In the course of using the Yule-Walker relations to conduct the fitting of ZNZ patterned VAR (p) models, as described in appendix 1, only the following p+1 lag covariance matrices are required to compute the estimated coefficient matrices and residual variance-covariance matrix::

$$\Gamma_0, \Gamma_1, \ldots, \Gamma_p \ .$$

However, the estimated V using the usual least squares (LS) method is as follows:

$$\hat{V} = \frac{1}{N-p} \sum_{i=p+1}^{N} \hat{\varepsilon}_i \hat{\varepsilon}_i' \ ,$$

where $\hat{\varepsilon}_i$ denotes the estimate of $\varepsilon(i)$.

This method suffers from the need to estimate and store all individual mx1 residual vectors, $\hat{\varepsilon}_t$, $t = 1, 2, \ldots, N$ and then compute $\hat{V}$ for each proposed ZNZ patterned VAR model. In order to estimate individual residual vectors, all observation vectors $u(t)$, $t = 1, 2, \ldots, N$ must be held in storage for conducting $\hat{\varepsilon}_t$ estimation. When using the LS method, a very large number of candidate ZNZ patterned VAR models must be estimated before the optimal model can be selected, which involves a considerable amount of computational cost in terms of execution time and memory storage and these costs are important considerations. Many

researchers, working with large samples will be aware of this inefficient procedure. It is obvious that estimation of the residual variance-covariance matrix, which minimizes the need for computing resources, becomes an important issue. As outlined later in the paper, there is no need to estimate individual residual vectors when an adjustment is made to the Yule-Walker approach. This approach is simple and avoids a considerable amount of computational costs.

The issue in Penm and Terrell (1984) is that their estimate of V using the Yule-Walker residual variance-covariance relation of (5) is not analyzed. Only the Yule-Walker coefficient relations in (4) are canvassed. A direct extension of the Yule-Walker residual variance-covariance relation to fit the ZNZ patterned VAR model is inappropriate as it is inconsistent with statistical theory. The problem is that the resultant estimated variance-covariance matrix of the white noise process becomes non-symmetric, violating the condition that V must be symmetric. This violation has important implications. One consequence is that VAR cannot be converted to an equivalent vector moving average (VMA) model as proposed in Penm and Terrell (1986) to conduct testing for Granger-causality. Further, innovation accounting proposed by Lee (1992) will not work under these conditions (Brailsford et al. 2001). Hence, this failure to ensure symmetry of estimates of V creates the motivation for developing an adjustment in this paper to the Yule-Walker relations for fitting of ZNZ patterned VAR models.

An alternative solution is to use other approaches that do not rely on the Yule-Walker relations. However, each of these approaches are problematic, particularly in terms of large computational costs. These alternative methods are briefly outlined below.

First, consider the standard least squares (LS) approach. As described in appendix 1, for fitting a full-order VAR (p) model using the Yule-Walker relations, the following block Toeplitz matrix C_{p+1} can be constructed:

$$C_{p+1} = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \cdots & \Gamma_p \\ \Gamma_{-1} & \Gamma_0 & & \Gamma_{p-1} \\ \vdots & \vdots & & \vdots \\ \Gamma_{-p} & \Gamma_{1-p} & \cdots & \Gamma_0 \end{bmatrix} = \begin{bmatrix} C_p & \vdots \\ \cdots & \Gamma_0 \end{bmatrix}, \qquad (6)$$

and the following relation for the estimate of $|V|$:

$$\left|\hat{V}_p\right| = \left|C_{p+1}\right| / \left|C_p\right| = \left|d_{p+1}\right|, \tag{7}$$

where d_i, $i = 1,..., p+1$ are diagonal block entries of the block diagonal matrix resulting from a block Choleski decomposition for $C_{p+1.}$ This outcome indicates that in the course of computing $\left|\hat{V}_p\right|$ for the VAR (p) model, the generalised residual sums of squares[3] for all the lower order VAR models fitted to the data are also obtained. However as described in appendix 2, this outcome cannot be achieved by using the conventional LS approach. Since $R_{LS}(p)$ for each different VAR model must be reconstructed from the observations to conduct individual fittings, and the observations must be saved in storage for reconstructing $R_{LS}(i)$, $i = 1,..., p$, a considerable increase in computational costs, based on execution time and data storage, will be required. Note that these weaknesses of the conventional LS method also exist in the remaining steps of selecting the optimal VAR, and become severe when the number of lags, or the number of variables, is large. Thus the commonly employed LS approach is considerably more computationally costly than the Yule-Walker approach.

Second, the generalized least squares (GLS) method can be conducted by applying the conventional LS approach as a basis. After the symmetric and positive definite $\hat{V}$ is estimated by the LS method, there exists an mxm non-singular matrix $\hat{K}$, such that $\hat{V}^{-1} = \hat{K}\hat{K}'$. We can pre-multiply $u(t)$ by $\hat{K}^{-1}$, and then follow the LS estimation for fitting of the VAR models to obtain the conventional GLS estimates. However as the LS approach to conduct the selection of the optimal ZNZ patterned VAR is computationally expensive when the number of possible candidate models could be billions, the conventional GLS method will similarly suffer from excessive computational costs.

Third, the maximum likelihood (ML) approach is a non-linear approach but becomes infeasible whenever the number of parameters is large (Chen and Zadrozny 1998). In addition there exist innumerable

3. The proposed model selection criteria use the generalised residual sums of squares

candidate models in the ZNZ patterned VAR environment. The ML approach needs to apply to each individual VAR model separately, and no previous computational information can be utilised.

Chen and Zadrozny (1998) propose the extended Yule-Walker equation to estimate a VAR for mixed frequency data. The estimated V for their approach is as follows:[4]

$$\hat{V} = \frac{1}{N-p} \sum_{i=p+1}^{N} \hat{\varepsilon}_i \hat{\varepsilon}_i',$$

which is identical to the conventional LS approach. Thus the approach of Chen and Zadrozny also needs to consider each VAR model independently for estimation of the individual residual variance-covariance matrices. In complete data cases (ie. no missing values), their approach only concerns full-order models. The ZNZ patterned modeling with no missing data is not investigated in Chen and Zadrozny.[5]

In appendix 2, it is shown that the conventional LS method is quite different from the Yule-Walker approach. Thus, $\hat{V}$ using the LS method is also quite different from $\hat{V}$ under the Yule-Walker approach. It follows that the approach of Chen and Zadrozny (1998) has ignored the issue of estimating the residual variance-covariance matrix. Although $\hat{V}$ using the LS method is asymptotically equivalent to $\hat{V}$ using the Yule-Walker approach, these two estimators can be quite different in a finite sample. If $\hat{V}$ proposed in Chen and Zadrozny is estimated using the Yule-Walker approach, then in the case of complete data the approach in the current paper can be employed to select the optimal ZNZ patterned VAR. Thus, again a considerable amount of computational costs can be avoided.

The most successful applications in ZNZ patterned VAR modeling are associated with Granger non-causality and indirect causality detection. This is because both Granger non-causality and indirect

4. See section 3 in Chen and Zadrozny (1998).

5. However the approach in Chen and Zadrozny (1998) addresses an interesting topic of estimation for mixed frequency data. Incorporating their approach into the ZNZ patterned modeling for mixed frequency data deserves further investigation.

causality detections are crucially dependent on making use of zero coefficient entries in the true structure, where the structure does indeed include several zero entries. Application of VARs to economic and financial time series data has revealed that zero entries are indeed possible (Caines et al. 1981, Penm et al. 1992). Since the ZNZ patterned VAR modeling allows for zero entries, the selected optimal ZNZ patterned VAR provides a straightforward and effective means of indicating all Granger-causality, Granger non-causality and indirect causality from the coefficient matrices on the lagged terms.

III. The Inconsistency in Using Yule-Walker Relations

In this section, the theoretical inconsistency of the use of the Yule-Walker relations for fitting of ZNZ patterned VAR models is demonstrated using a two-asset example.

In considering a ZNZ patterned VAR model, zero entries in the parameter matrices A_k of (1) are allowed. If $y_{1,t}$ and $y_{2,t}$ are the log prices of the assets, then the returns on the assets are defined by $\Delta y_{1,t} = z_{1,t}$ and $\Delta y_{2,t} = z_{2,t}$. Both $z_{1,t}$ and $z_{2,t}$ are jointly determined by the following two equations:

$$z_{1,t} + a_{12} z_{2,t-1} = \varepsilon_{1,t} , \tag{8}$$

$$z_{2,t} + a_{21} z_{1,t-1} + a_{22} z_{2,t-1} = \varepsilon_{2,t} . \tag{9}$$

In this two-equation system the first equation shows that $z_{1,t}$ is caused by $z_{2,t}$, while the second equation indicates that $z_{2,t}$ is caused by $z_{1,t}$, thereby creating a feedback relation.

The equivalent VAR model of this system can then be expressed as:

$$\begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} 0 & a_{12} \\ a_{21} & a22 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \tag{10}$$

where the white noise process comprises two components $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$,

with:

$$E\left\{\varepsilon_{1,t}\right\} = E\left\{\varepsilon_{2,t}\right\} = 0\,,$$

and

$$E\left\{\begin{bmatrix}\varepsilon_{1,t}\\ \varepsilon_{2,t}\end{bmatrix}\begin{bmatrix}\varepsilon_{1,t-k} & \varepsilon_{2,t-k}\end{bmatrix}\right\}\begin{matrix}=V \text{ as } k=0\\ =0 \text{ as } k>0\end{matrix}.$$

The Yule-Walker coefficient relations are now used to estimate a_{12}, a_{21}, and a_{22}. Since $a_{11} = 0$, and $\varepsilon_{1,t}$ is uncorrelated with the asset return, $z_{2,t-1}$, the following relation is apparent:

$$E\left\{z_{1,t}z_{2,t-1}\right\} + \hat{a}_{12}E\left\{z_{2,t-1}z_{2,t-1}\right\} = 0\,,$$

where $\hat{a}_{ij}$ are the estimates of a_{ij}. Thus we will achieve $\hat{a}_{12} = -\tau_{12}(1)/\tau_{22}(0)$, where the correlation functions between asset returns,

$$= \tau_{ji}(k) = \frac{1}{N}\sum_{t=1}^{N-k} z_i(t+k)z_j(t) = \tau_{ij}(-k).$$

From (10), since the asset return vector $\begin{bmatrix} z_{1,t-1} & z_{2,t-1}\end{bmatrix}'$ is uncorrelated with $\begin{bmatrix}\varepsilon_{1,t} & \varepsilon_{2,t}\end{bmatrix}'$, the following arises:

$$\begin{bmatrix}\tau_{11}(0) & \tau_{21}(0)\\ \tau_{12}(0) & \tau_{22}(0)\end{bmatrix}\begin{bmatrix}\hat{a}_{21}\\ \hat{a}_{22}\end{bmatrix} = -\begin{bmatrix}\tau_{21}(1)\\ \tau_{22}(1)\end{bmatrix}. \tag{11}$$

Hence,

$$\hat{a}_{21} = \frac{\tau_{21}(0)\tau_{22}(1) - \tau_{22}(0)\tau_{21}(1)}{\tau_{11}(0)\tau_{22}(0) - \tau_{21}(0)\tau_{12}(0)}\,,$$

and

$$\hat{a}_{22} = \frac{\tau_{12}(0)\tau_{21}(1) - \tau_{11}(0)\tau_{22}(1)}{\tau_{11}(0)\tau_{22}(0) - \tau_{21}(0)\tau_{12}(0)}.$$

Thus the coefficient estimates in terms of the correlation functions between asset returns are established. Of note, the use of the above approach is identical to the use of equation (A.1.3) as proposed in appendix 1 for fitting of the ZNZ patterned VAR models.

As a result the estimate $\hat{V}$ in equation 5 becomes:

$$\hat{V} = \begin{bmatrix} \tau_{11}(0) & \tau_{12}(0) \\ \tau_{12}(0) & \tau_{22}(0) \end{bmatrix}$$

$$+ \begin{bmatrix} \hat{a}_{12}\tau_{12}(1) & \hat{a}_{12}\tau_{22}(1) \\ \hat{a}_{21}\tau_{11}(1) + \hat{a}_{22}\tau_{12}(1) & \hat{a}_{21}\tau_{21}(1) + \hat{a}_{22}\tau_{22}(1) \end{bmatrix}, \qquad (12)$$

which is non-symmetric. Intuitively, V is symmetric in the true model of (1) and there is a need for the estimate $\hat{V}$ to conform to the behaviour of V. Therefore the estimate $\hat{V}$ must be a symmetric matrix. As described earlier, this non-symmetric $\hat{V}$ violates the symmetric condition required in Lee (1992) and in Penm and Terrell (1986). This violation indicates that, in practice, the innovation accounting described in Lee will not work (Brailsford et al. 2001), and a VAR model cannot be converted to its equivalent VMA model as proposed in Penm and Terrell to conduct testing for Granger-causality. Thus an adjustment to the Yule-Walker relations is required.

IV. The Adjustment

The necessary adjustment to the Yule-Walker relations for fitting of VAR models with ZNZ patterned coefficient matrices follows directly from the inconsistency demonstrated in the previous section.

With the definition of the variance-covariance matrix in (2),

$$V = E\left\{\begin{bmatrix}\varepsilon_{1,t}\\ \varepsilon_{2,t}\end{bmatrix}\begin{bmatrix}\varepsilon_{1,t} & \varepsilon_{2,t}\end{bmatrix}\right\}$$

$$= E\left\{\begin{bmatrix}z_{1,t}\\ z_{2,t}\end{bmatrix}+\begin{bmatrix}0 & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}z_{1,t-1}\\ z_{2,t-1}\end{bmatrix}\right\}$$

$$\left\{\begin{bmatrix}z_{1,t} & z_{2,t}\end{bmatrix}+\begin{bmatrix}z_{1,t-1} & z_{2,t-1}\end{bmatrix}\begin{bmatrix}0 & a_{21}\\ a_{12} & a_{22}\end{bmatrix}\right\}$$

Then the estimate

$$\hat{V} = \begin{bmatrix}\tau_{11}(0) & \tau_{12}(0)\\ \tau_{12}(0) & \tau_{22}(0)\end{bmatrix}$$

$$+\begin{bmatrix}\hat{a}_{12}\tau_{12}(1) & \hat{a}_{12}\tau_{22}(1)\\ \hat{a}_{21}\tau_{11}(1)+\hat{a}_{22}\tau_{12}(1) & \hat{a}_{21}\tau_{21}(1)+\hat{a}_{22}\tau_{22}(1)\end{bmatrix}$$

$$+\begin{bmatrix}\hat{a}_{12}\tau_{12}(1) & \hat{a}_{21}\tau_{11}(1)+\hat{a}_{22}\tau_{12}(1)\\ \hat{a}_{12}\tau_{22}(1) & \hat{a}_{21}\tau_{21}(1)+\hat{a}_{22}\tau_{22}(1)\end{bmatrix} \tag{13}$$

$$+\begin{bmatrix}0 & \hat{a}_{12}\\ \hat{a}_{21} & \hat{a}_{22}\end{bmatrix}\begin{bmatrix}\tau_{11}(0) & \tau_{12}(0)\\ \tau_{12}(0) & \tau_{22}(0)\end{bmatrix}\begin{bmatrix}0 & \hat{a}_{21}\\ \hat{a}_{12} & \hat{a}_{22}\end{bmatrix}$$

Since the first matrix of equation 13 is symmetric, the second matrix is the transpose of the third matrix, and the remaining product matrix is also symmetric, therefore the matrix $\hat{V}$ is symmetric.

An analogous approach using equation 13 is feasible. From equation 2,

$$V = E\left\{\left[\sum_{k=0}^{p} A_k u(t-k)\right]\left[\sum_{j=0}^{p} u'(t-j)A_j'\right]\right\}.$$

Then,

$$\hat{V} = \Gamma_0 + \sum_{k=1}^{p} \hat{A}_k \Gamma_{-k} + \sum_{j=1}^{p} \Gamma_j \hat{A}_j' + \sum_{j=1}^{p} \sum_{k=1}^{p} \hat{A}_k \Gamma_{j-k} \hat{A}_j'. \tag{14}$$

It is obvious that this $\hat{V}$ is symmetric. Since $\Gamma_{-j}' = \Gamma_j$, Γ_0 is symmetric. If k is redefined as j, the second matrix is the transpose of the third matrix; and if j and k are redefined as k and j respectively, the fourth matrix becomes

$$\sum_{k=1}^{p} \sum_{j=1}^{p} \hat{A}_j \Gamma_{k-j} \hat{A}_k' ,$$

which is the transpose of the fourth matrix itself.

In addition, comparing equation 14 to the estimator of V using individual residual vectors, the structure of (14) is computationally efficient in terms of both execution time and storage requirements, and provides the obvious relations to link the covariance matrices with different lags.

Of note, consideration of the contemporaneous correlation in $\varepsilon(t)$ cannot be ignored. A ZNZ patterned VAR model can be viewed as a system of 'seemingly unrelated regressions' as originally proposed by Zeller (1962). As the regressors in each equation of the VAR model are no longer necessarily the same, the generalised least squares (GLS) coefficient estimator using the Yule-Walker relations for the ZNZ patterned VAR is more efficient than the estimator using equation (14). Brailsford et al. (2001) show that this GLS estimator is an approximation to the ML estimator asymptotically. Henceforth the notation GLS-YW is used for this estimator. As described earlier, $\hat{V}^{-1} = \hat{K}\hat{K}'$. We premultiply $u(t)$ by $\hat{K}^{-1}$, and then follow the proposed method of using the Yule-Walker relations for fitting of VAR models, and so obtain the GLS-YW coefficient estimates of the ZNZ patterned VAR model.

V. Empirical Testing

In this section, an application is presented to illustrate the practical use of the algorithm. First, consider a potential causality relationship between

stock and exchange rate markets. For instance, flows of capital influence exchange rate movements and such flows have been shown to be related to equity returns (eg. Froot et al. 1998). Of course, the relationship between exchange rates and stock prices is more complex than implied here and involves consideration of parity conditions and inflationary expectations.[6] Nevertheless, the purpose here is illustrative only and shows how the effect of correctly estimating the residual variance-covariance matrix can provide insights into Granger-causal relationships among financial variables. In a similar vein, Bekaert and Hodrick (1992) examine the predictability of excess returns on equity and foreign exchange markets using a six-variable VAR model.

The sample period is chosen as 1 January 1999 to 31 December 1999 and all data are sampled daily. The selected equity market is Hong Kong. This market is an international financial centre, and allows free flows of international funds. Within this context, the following three variables are studied contemporaneously in a stochastic vector system using the ZNZ patterned VAR modeling:

(i) Euro to US dollar - exchange rate (EUFX)
(ii) Hong Kong's Hang Seng - stock price index (HSI)
(iii) Hong Kong dollar to US dollar - exchange rate (HKFX)

The Hang Seng Index is the main stock market indicator in Hong Kong. This index comprises 33 constituent stocks which are representative of the market. The aggregate market capitalization of these stocks accounts for about 70% of the total market capitalization on Hong Kong's stock exchange. At the beginning of 1999 the HSI was 9,000. However it climbed to 17,000 by the end of 1999, closing with a 90% gain over the year.

The Hong Kong Monetary Authority uses a pegged exchange rate fixed at HK$7.8 to US$1. However, deviations from this rate occur through the dealer market.[7] The sample period coincides with the introduction of the Euro.

---

6. The relationship between stock prices and exchange rates has a considerable literature which involves purchasing power parity. For discussion, see Frenkel (1981), Dumas and Solnik (1995).

7. In Hong Kong, only the Hong Kong Monetary Authority uses the fixed exchange rate. This fixed rate does not apply to other dealers and fluctuations thus occur.

TABLE 1. Orthogonal Polynomial Regression for Trend Removal

	Intercept	Orthogonal polynomial P_1
Log(EUFX)	–.1159	.0004
	(.0028)	(.0000)
Log(HSI)	9.2064	.0018
	(.0093)	(.0000)
Log(HSFX)	2.0469	.0001
	(.0000)	(.0000)

Note: This table reports the results of fitting a first-order polynomial to the three time series under examination. The series comprise daily data sampled over the period 1/1/1999 to 31/12/1999 and are the Euro (EUFX), Hang Seng Stock Index (HIS) and the Hong Kong dollar exchange rate (HSFX). The values in parentheses are standard errors of the coefficient estimates.

The variables are log transformed such that $u_1(t)$=log(EUFX), $u_2(t)$=log(HSI) and $u_3(t)$=log(HKFX). Following Penm and Terrell (1984), Forsythe's (1957) method is initially used for generating orthogonal polynomials to assess the data for suitable detrending to produce stationarity.[8] The results show that detrending using a first-order polynomial is required before fitting the VAR models. The coefficient estimates and associated standard errors from fitting the polynomial are reported in table 1.

After detrending, a maximum order of 36 is assigned and the search procedures proposed in Penm and Terrell (1984) are employed to obtain the optimal ZNZ patterned VAR model. Each of three order selection criteria - Akaike, Schwarz and Hannan - is used to determine the best specification. The ability of these three order selection criteria to determine the true specification of a stationary VAR has been examined using a simulation approach by Penm and Terrell (1984). Their results suggest that the Schwarz criterion (*SC*) is superior in order-identification to the other two alternatives in ZNZ patterned VAR modeling for causality studies. Therefore, the specification determined by *SC* is selected and used as the benchmark model for analysing lead-lag relations.

8. There is a large literature that traverses the best method of removing non-stationarity in a series. The approach here involves a polynomial trend, because of its ability to remove the long-term non-stationarity in any series without seriously impacting on other cyclical variation closely adjacent to non-stationary long-term movements.

TABLE 2. The Optimal ZNZ Patterned VAR

Coefficient Estimator	The Yule-Walker Approach	GLS-YW
Type of coefficient matrices selected	$\hat{A}_1 \begin{bmatrix} -.9630 & 0 & 0 \\ (.0172) & & \\ -.1680 & -.9416 & 0 \\ (.0585) & (.0183) & \\ .0009 & 0 & -.8700 \\ (.0004) & & (.0311) \end{bmatrix}$	$\begin{bmatrix} -.9630 & 0 & 0 \\ (.0171) & & \\ -.0061 & -.9409 & 0 \\ (.0016) & (.0172) & \\ .0053 & 0 & -.8705 \\ (.0018) & & (.0303) \end{bmatrix}$
Estimate of Residual variance-covariance matrix (x 10^{-4})	$\begin{bmatrix} .3783 & .1239 & -.0007 \\ .1239 & 2.732 & .0006 \\ -.0007 & .0006 & .0001 \end{bmatrix}$	

Residual analysis	p	0	1	2	3	4	5
Value of *SC*		1.0	1.031	1.065	1.102	1.141	1.180

Note: This table presents the results from fitting a VAR to the three time series under examination. The series comprise daily data sampled over the period 1/1/1999 to 31/12/1999 and are the Euro (EUFX), Hang Seng Stock Index (HIS) and the Hong Kong dollar exchange rate (HSFX). The values in parentheses are standard errors of the non-zero coefficient estimates. The selection criterion employed is Schwarz criterion (*SC*) and normalised values of *SC* are presented. The optimal lag selected for the variables under investigation in the VAR is 1. The residual analysis confirms the residuals have white noise characteristics.

$$u(t) = \{\log \text{EUFX}, \log \text{HSI}, \log \text{HKFX}\}'$$

The coefficient estimates of the chosen specification using the adjusted Yule-Walker relations are presented in table 2. To check the adequacy of the model fit, the strategy suggested in Tiao and Tsay (1989) is used, with the proposed algorithm applied to test the residual vector series, using the *SC* criterion.[9] The results in table 2 support the residual vector being a white noise process. The procedures outlined in section IV to obtain the GLS-YW estimator are then carried out, with

9. Tiao and Tsay (1989) proposed an algorithm using the crit(m,j) criterion to select the vector autoregressive moving average process with zero entries. After the final model is selected, their algorithm was then applied to the residual series to test whether this series is a vector white noise process.

the resultant comparative output also presented in table 2.

The relationships identified by the three selection criteria are markedly similar. All the determined specifications consistently indicate that the Euro is the major variable which provides leading information for the other components of the system. The lagged Euro enters not only its own equation but also those of HSI and HKFX. In all the determined specifications, the lagged level of HSI does not enter any of the exchange rate equations, indicating that variations in the Hong Kong stock market index provide little leading information for the exchange rate markets, as expected. Also no lagged HKFX components enter the equation of the HSI and EUFX, indicating that this variable contains little leading information for either the local stock market or the Euro. Given that the HKFX is relatively stable, this latter result is not surprising.

The more surprising result is the influence of the Euro. Hong Kong is known as an open market and there are considerable international capital flows into and out of its stock market. A major component of these flows is from Europe. Moreover, it is well-known that the weakness of the Euro has been a salient feature in international markets over the past two years. As noted above, over the sample period, the Hong Kong market surged by 90%. Hence, a link between the Euro and the Hong Kong stock market is feasible.

A more complete analysis would include other economic and financial variables such as net capital flows, interest rates and money supply, which could all play a significant role. Indeed, our model could be extended to incorporate the recent work of Bekaert et al. (1999) who propose a larger system. However, as discussed earlier, this empirical examination is for illustrative purposes. The importance of this application is that it shows the procedures which can be applied to any set of variables. As an example, in the context of emerging markets where traditional models and theories have met with little success, such exercises are likely to provide valuable insights into the relationships and causality between financial variables.

VI. Conclusion

The use of VAR modeling in financial economics has become common. However, the models are typically constrained through problems of

over-parameterisation. In this paper, an adjustment to the Yule-Walker relations for fitting of ZNZ patterned VAR models is presented. The adjustment is consistent with statistical procedures in theory and has the advantages of computational efficiency and reliability.

The procedure has been applied to the Hong Kong stock market, focusing on its relationship with international foreign exchange markets. The results of this exercise are helpful in understanding linkages between various markets and/or financial variables. As indicated above, in the area of emerging markets where there is often no clear consensus concerning relationships among financial variables and each market appears to exhibit almost unique characteristics, this procedure can potentially yield important insights.

Appendix 1: The Use of the Yule-Walker Relations for Fitting VAR Models

Three model selection criteria are employed to select the optimal ZNZ patterned VAR. They are:

$$AIC = \log\left|\hat{V}_p\right| + [2/N]S \ ,$$

$$HC = \log\left|\hat{V}_p\right| + [2\log\log N/N]S \ ,$$

$$SC = \log\left|\hat{V}_p\right| + [\log N/N]S \ ,$$

where S is the number of functionally independent parameters estimated.

The detailed method of selecting the optimal ZNZ patterned VAR with the smallest value of each selection criterion is summarised in Brailsford et al. (2001).

A. Fitting of full-order VAR models

In considering the use of the Yule-Walker coefficient relations for fitting of a full-order VAR(p) model of (1), equation 4 can be expressed as:

$$\Lambda_p R_p = -\Pi_p \,, \tag{A.1.1}$$

where

$$\Lambda_p = \left\{ \hat{A}_1 \quad \hat{A}_2 \quad \dots \quad \hat{A}_p \right\},$$

$$\Pi_p = \left\{ \Gamma_1 \quad \Gamma_2 \quad \dots \quad \Gamma_p \right\},$$

and

$$R_p = \begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \cdots & \Gamma_{1-p} \\ \Gamma_1 & \Gamma_0 & & \Gamma_{2-p} \\ \vdots & \vdots & & \vdots \\ \Gamma_{p-1} & \Gamma_{p-2} & \cdots & \Gamma_0 \end{bmatrix}.$$

Analogously, to fit a VAR (p+1) model, we have:

$$R_{p+1} = \begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \cdots & \Gamma_{-p} \\ \Gamma_1 & \Gamma_0 & & \Gamma_{1-p} \\ \vdots & \vdots & & \vdots \\ \Gamma_p & \Gamma_{p-1} & \cdots & \Gamma_0 \end{bmatrix} = \begin{bmatrix} R_p & \vdots \\ \cdots & \Gamma_0 \end{bmatrix}.$$

Now form a block Toeplitz matrix C_{p+1} of (6), and have the following block Choleski decomposition:

$$C_{p+1} = L_{p+1} D_{p+1} L'_{p+1} \,, \tag{A.1.2}$$

where L_{p+1} is a lower block triangular matrix, and D_{p+1} is a block diagonal matrix with diagonal block entries d_i, i = 1,..., p+1.

Thus, we have $\left|C_{p+1}\right| = \left|C_p\right| \left|\Gamma_0 - \Gamma_p R_p^{-1} \Gamma'_p\right| = \left|C_p\right| \left|\hat{V}_p\right|$, and can achieve the outcome: $\left|\hat{V}_p\right| = \left|d_{p+1}\right|$. More importantly, in the course of computing $\left|\hat{V}_p\right|$, d_i, $i = 1,\dots,p+1$ will be obtained by using (A.1.2). Since $\left|d_i\right|$ is the $\left|\hat{V}\right|$ for the VAR (i–1) model, the generalised residual sum of squares for all the lower order VAR models fitted to the data are also obtained.

Therefore a considerable amount of computational cost can be avoided.

B. Fitting of ZNZ patterned VAR models

In considering the use of the Yule-Walker coefficient relations for fitting of ZNZ patterned VAR models of (1), the coefficient estimates obey the following relationship:

$$Z_p(C_r)\alpha(C_r) = \gamma(C_r), \tag{A.1.3}$$

where $Z_p = \{I_m \otimes R_p\}$, $\alpha = vec\{\Lambda'_p\}$, $\gamma = vec\{\Pi'_p\}$, C_r is an integer set which contains $c_1, c_2, \ldots, c_r$, and the $(c_1, c_2, \ldots, c_r)$ th entries of α are constrained to zero. Then $\alpha(C_r)$ and $\gamma(C_r)$ are formed by placing 0 in the $(c_1, c_2, \ldots, c_r)$ the row entries of α and γ, and $Z(C_r)$ is formed by placing 1 in the $\{(c_1, c_1), (c_2, c_2), \ldots, (c_r, c_r)\}$ diagonal entries of Z and 0 everywhere else in the $(c_1, c_2, \ldots, c_r)$ rows and columns of Z.
Also, the estimate of V is:

$$\hat{V} = \Gamma_0 + \sum_{k=1}^{p} \hat{A}_k \Gamma_{-k} + \sum_{j=1}^{p} \Gamma_j \hat{A}'_j + \sum_{j=1}^{p} \sum_{k=1}^{p} \hat{A}_k \Gamma_{j-k} \hat{A}'_j ,$$

which is the equation (14) in this paper.

Note that only the p+1 lag covariance matrices shown in (3) are required to compute $\hat{A}_k$ and $\hat{V}$ by using the Yule-Walker approach. Further, fitting of subset VAR models, which are the VAR models with intermediate lags constrained to zero matrices, can also be achieved by using both (A.1.3) and (14).

Appendix 2: The Use of the LS Method for Fitting of VAR Models

To fit a full order VAR (p) model of (1) for a given set of observations $\{u(t), t = 1,..., N\}$, the estimated V using the usual least squares (LS)

method is as follows:

$$\hat{V} = \frac{1}{N-p} \sum_{i=p+1}^{N} \hat{\varepsilon}_i \hat{\varepsilon}_i' , \tag{A.2.1}$$

where $\hat{\varepsilon}_i$ denotes the estimate of $\varepsilon(i)$.

For simplicity, consider the scalar case. Thus m=1. Equation (A.2.1) can be rewritten as:

$$\hat{V} = \frac{1}{N-p} \sum_{i=p+1}^{N} \left[u(t) - \sum_{k=1}^{p} \hat{a}_k u(t-k) \right] \left[y(t) - \sum_{k=1}^{p} \hat{a}_k u(t-k) \right].$$

Thus the associated linear regression model can be expressed as:

$$\begin{bmatrix} u(N) \\ \vdots \\ u(p+1) \end{bmatrix} = \begin{bmatrix} u(N-1) & \cdots & u(N-p) \\ \vdots & \vdots & \vdots \\ u(p) & \cdots & u(1) \end{bmatrix} \begin{bmatrix} -a_1 \\ \vdots \\ -a_p \end{bmatrix} + \begin{bmatrix} \varepsilon(N) \\ \vdots \\ \varepsilon(p+1) \end{bmatrix}.$$

The usual least squares estimate of $\beta = \begin{bmatrix} -a_1 & \cdots & -a_p \end{bmatrix}$ in the model

$$Y = X\beta + \eta ,$$

is then:

$$\hat{\beta} = (X'X)^{-1} X'Y .$$

Thus we have:

$$R_{LS}(p) = (X'X)$$

$$= \begin{bmatrix} \sum_{i=p}^{N-1} u^2(i) & \cdots & \sum_{i=p}^{N-1} u(i)u(i-p+1) \\ \vdots & \vdots & \vdots \\ \sum_{i=p}^{N-1} u(i)u(i-p+1) & \cdots & \sum_{i=1}^{N-1} u^2(i) \end{bmatrix}.$$

Analogously, to fit a VAR (p–1) model, we have:

$$R_{LS}(p-1)=$$

$$=\begin{bmatrix} \sum_{i=p-1}^{N-1} u^2(i) & \cdots & \sum_{i=p-1}^{N-1} u(i)u(i-p+1) \\ \vdots & \vdots & \vdots \\ \sum_{i=p-1}^{N-1} u(i)u(i-p+1) & \cdots & \sum_{i=1}^{N-1} u^2(i) \end{bmatrix}.$$

However,

$$R_{LS}(p) \neq \begin{bmatrix} R_{LS}(p-1) & \vdots \\ \cdots & . \end{bmatrix}.$$

Note that every (i,j) entry of $R_{LS}(p-1)$ is different from the (i,j) entry of $R_{LS}(p)$. It is obvious that the usual LS method is quite different from the Yule-Walker approach.

Thus, in fitting a VAR (p) model, the generalised residual sum of squares for all the lower order VAR models fitted to the data cannot be obtained by using the LS method. Note that these weaknesses of the LS method also exist in fitting for VAR models with zero entries, and become severe when the number of lags of (1), or the number of variables in the system, is large.

References

Bekaert, G.; Harvey, C. R.; and Lumsdaine, R.L. 1999. The dynamics of emerging market equity flows. Working paper. National Bureau of Economic Research.

Bekaert, G., and Hodrick, R. J. 1992. Characterizing predictable components in excess returns on equity and foreign exchange markets. *Journal of Finance* 47:467-509.

Bhattacharya, U.; Daouk, H.; Jorgenson, B.; and Kehr, C. 2000. When an event is not an event: The curious case of an emerging market. *Journal of Financial Economics* 55:69-101.

Brailsford, T. J.; Penm, J. H. W.; and Terrell, R. D. 2001. Supplementary mimeo to the adjustment of the Yule-Walker relations in VAR modeling: The impact of the Euro on the Hong Kong stock market. Unpublished paper. Australian National University.

Caines, P. E.; Keng, C. W.; and Sethi, S. P. 1981. Causality analysis and multivariate autoregressive modeling with an application to supermarket sales analysis. *Journal of Economic Dynamics and Control* 3:267-298.

Chen, B., and Zadrozny, P. A. 1998. An extended Yule-Walker method for estimating a vector autoregressive model with mixed-frequency data. *Advances in Econometrics* 13: 47-73.

Dumas, B., and Solnik, B. 1995. The world price of foreign exchange risk. *Journal of Finance* 50: 445-479.

Eun, C., and Shim, S. 1989. International transmission of stock market movements. *Journal of Financial and Quantitative Analysis* 24: 241-256.

Forsythe, G. E. 1957. Generation and use of orthogonal polynomials for fitting data with a digital computer. *SIAM Journal on Applied Mathematics* 5: 74-88.

Frenkel, J.A. 1981. The collapse of purchasing power parities during the 1970s. *European Economic Review* 16: 145-166.

Froot, K. A.; O'Connell, P.; and Seasholes, M. 1998. The portfolio flows of international investors. Working paper. Harvard Business School.

Granger, C. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37: 424-438.

Hassapis, C.; Pittis, N.; and Prodromidis, K. 1999. Unit roots and Granger causality in the EMS interest rates: The German dominance hypothesis revisited. *Journal of International Money and Finance* 18: 47-73.

Lee, B. S. 1992. Causal relations among stock returns, interest rates, real activity, and inflation. *Journal of Finance* 47: 1591-1603.

Malliaris, A. G., and Urrutia, J. L. 1992. The international crash of October 1987: Causality tests. *Journal of Financial and Quantitative Analysis* 27: 353-364.

Penm, J. H. W.; Penm, J. H. C.; and Terrell, R. D. 1992. Using the bootstrap as an aid in choosing the approximate representation for vector time series. *Journal of Business and Economic Statistics:* 213-219.

Penm, J. H. W., and Terrell, R. D. 1984. Multivariate subset autoregressive modeling with zero constraints for detecting causality. *Journal of Econometrics* 24: 311-330.

Penm, J. H. W., and Terrell, R. D. 1986. The 'derived' moving average model and its role in causality. *Journal of Applied Probability* 23: 99-111.

Thorbecke, W. 1997. On stock market returns and monetary policy. *Journal of Finance* 52: 635-654.

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

Whitelaw, R. F. 1994. Time variations and covariations in the expectation and volatility of stock market returns. *Journal of Finance* 49: 515-541.

Zeller, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57: 348-368.