Electoral Studies 31 (2012) 852–860

Contents lists available at SciVerse ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Maximum likelihood estimation of structural VARFIMA models

Wen-Jen Tsay*,[1]

The Institute of Economics, Academia Sinica, 128 Academia Road, Sec. 2, Taipei, Taiwan, ROC

ARTICLE INFO

Article history:
Received 13 January 2012
Received in revised form 31 May 2012
Accepted 22 June 2012

Keywords:
Fractional cointegration
Structural VARFIMA
Durbin–Levinson algorithm

ABSTRACT

This paper considers the maximum likelihood estimation of a class of structural vector autoregressive fractionally integrated moving-average (VARFIMA) models. The structural VARFIMA model includes the fractional cointegration model as one of its special cases. We show that the conditional likelihood Durbin–Levinson (CLDL) algorithm of Tsay (2010a) is a fast and reliable approach to estimate the long-run effects as well as the short- and long-term dynamics of a structural VARFIMA process simultaneously. In particular, the computational cost of the CLDL algorithm is much lower than that proposed in Sowell (1989) and Dueker and Startz (1998). We apply the CLDL method to the Congressional approval data of Durr et al. (1997) and find that the long-run effect of economic expectations on Congressional approval is at least 0.5718, which is over twice the estimate of 0.24 found in Table 2 of Box-Steffensmeier and Tomlinson (2000). This paper also tests the divided party government hypothesis with the CLDL algorithm.

1. Introduction

This paper considers the maximum likelihood estimation of a class of structural vector autoregressive fractionally integrated moving-average VARFIMA (p,d,q), models. This model includes the fractional cointegration model as one of its special cases and has been investigated by Sowell (1989) and Dueker and Startz (1998). It also encompasses the stationary and invertible VARFIMA processes of Tsay (2010a) whereby all the data-generating processes (DGP) behind the structural VARFIMA models are stationary. The broad coverage of the structural VARFIMA model identifies itself as a useful workhorse for many political time series observations, including Box-Steffensmeier and Tomlinson (2000) and Clarke and Lebo (2003). In particular, in Section 4 of this paper we estimate the long-run effect of economic expectations on Congressional approval based on a 2-dimensional fractional cointegration model of Congressional approval data of Durr et al. (1997, DGW hereafter). The magnitude of the long-run effect should be the focus of the literature, because it signifies whether the economic prospects of the public strongly affect their support of the Congress. The use of the structural VARFIMA model allows us to simultaneously address the long-run effects as well as the short- and long-term dynamics characterized by the AR, MA, and the fractional differencing parameters.

This paper is strongly motivated by the observations in Box-Steffensmeier and Tomlinson (2000, p. 71) that the program of Dueker and Startz (1998) is extremely sensitive to the starting values and that their computation with Congressional approval data of DGW (1997) tends to get stuck in local minima. Another shortcoming of Sowell's (1989) algorithm is its heavy computational burden. Dueker and Startz (1998, p. 423) demonstrate that it takes 35 min on a 200-MHz PC for each iteration of the maximum likelihood estimation of a bivariate VARFIMA process with 121 observations and 18 parameters when implementing Sowell's (1989) algorithm.

* Tel.: +886 2 27822791x296; fax: +886 2 27853946.
E-mail address: wtsay@econ.sinica.edu.tw.

[1] The author thanks the Editor, Professor Clarke, and two anonymous referees for useful comments and suggestions. Part of the research was conducted when Tsay was a visiting scholar in the Department of Economics, University of Washington. The author would like to thank the financial and managerial support from the Fulbright Program. Research assistance from Shih-Tang Hwu is gratefully appreciated.

0261-3794/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.electstud.2012.06.007

This paper explains how the conditional likelihood Durbin–Levinson (CLDL) algorithm of Tsay (2010a) can be helpful to estimate the structural VARFIMA model efficiently. First, the CLDL algorithm is a one-step likelihood-based estimator and can estimate the aforementioned long-run effects as well as the short- and long-term dynamics of a structural VARFIMA process simultaneously. Because a one-step procedure is usually more efficient than a two-step or multiple-step procedure and is a consistent estimator if the model is correctly specified, the long-run effect estimate from the CLDL algorithm will generally be different from the one generated from Box-Steffensmeier and Tomlinson (2000) who employ a two-step procedure. Indeed, when applying the CLDL algorithm to test the divided party government hypothesis with the 80 observations of Congressional approval data of DGW (1997) based on a 3-dimensional VARFIMA (*p*,*d*,*q*) model in the following Section 5, we establish the first evidence that the disturbance term of the fractional cointegration model of Congressional approval might be a fractionally integrated process. This indicates that a flexible long memory process is required to capture the dynamic behavior of the data of Congressional approval and justifies the use of the structural VARFIMA model for this important issue of Political Science.

Second, the CLDL algorithm can evaluate the conditional likelihood function of the structural VARFIMA models exactly. This is in sharp contrast with the algorithm of Sowell (1989) and that of Dueker and Startz (1998), which are subject to a truncating error when the AR parameters are present.

The third advantage of the CLDL algorithm is that its computation is much faster than that proposed in Sowell (1989) and Dueker and Startz (1998), because it utilizes an efficient Durin–Levinson algorithm. Due to the high speed of computation, Tsay (2010a) conducts a Monte Carlo experiment to show the finite sample performance of the CLDL algorithm for 3-dimensional VARFIMA processes under a sample size of up to 400. Therefore, the use of the CLDL algorithm also resolves the comment of Lebo et al. (2000, p. 38) that "the only complaints about full maximum likelihood estimation concern its computationally intensive algorithm."

The remaining parts of this paper are arranged as follows: Section 2 presents the structural VARFIMA (*p*,*d*,*q*) models. Section 3 explains the implementation of the CLDL algorithm for the structural VARFIMA process. We apply the CLDL methodology to the data of DGW (1997) in Section 4. The major task is to estimate the long-run effect of economic expectations on Congressional approval using various 2-dimensional VARFIMA models. Section 5 tests the divided party government hypothesis with a 3-dimensional structural VARFIMA model. Section 6 provides a conclusion.

2. Structural VARFIMA models

Consider the structural multivariate time series model with fractionally integrated errors:

$$\begin{cases} y_t = \alpha_1^\top D_t + \beta^\top X_t + u_t, \\ X_t = \alpha_2 D_t + \gamma^\top X_{t-1} + V_t, \end{cases} \quad or \quad \begin{bmatrix} y_t - \alpha_1^\top D_t - \beta^\top X_t \\ X_t - \alpha_2 D_t - \gamma^\top X_{t-1} \end{bmatrix} = \begin{bmatrix} u_t \\ V_t \end{bmatrix} = W_t, \tag{1}$$

where D_t is a vector of deterministic functions, including a constant or linear trend, α_1 is a vector of parameters, and α_2 is a matrix of parameters conformable to D_t. Here, u_t is a univariate fractionally integrated process of order d_1, and V_t is an $r-1$ dimensional time series with $r-1$ potentially different orders of fractional integratedness. This model has been considered by Baillie and Bollerslev (1994), Cheung and Lai (1993), and Dueker and Startz (1998). When $r=2$ and $\gamma = 1$, we can easily see the well-known fractional cointegration model belongs to one of its special cases.

The model in eq. (1) is considered in Sowell (1989) and later employed by Dueker and Startz (1998) to describe the joint behavior of U.S. and Canadian bond rates. Essentially, the idea behind the algorithm of Sowell (1989) and Dueker and Startz (1998) for the model in eq. (1) is to follow the spirit of the conditional likelihood function in Box and Jenkins (1976, Chapter 7). In other words, conditional on the structural parameters $\{\alpha_1, \alpha_2, \beta, \gamma\} = \Psi$, we can write down the likelihood of W_t, provided the probability density function of W_t is known. The associated conditional likelihood is denoted as $L(W_t|\Psi)$. This notation signifies the likelihood function is computed conditional on the value of Ψ.

In the literature starting with Sowell (1992), W_t is usually assumed to be generated as:

$$\Phi(B) diag(\nabla^d) W_t = \Theta(B) Z_t, \tag{2}$$

where $t = 1, 2, \ldots, T$, W_t is an r-dimensional vector of observations of interest, and $\Phi(B)$ and $\Theta(B)$ are finite order matrix polynomials in B (usual lag operator), such that:

$$\Phi(B) = \Phi_0 - \Phi_1 B - \ldots - \Phi_p B^p, \quad \Theta(B) = \Theta_0 + \Theta_1 B + \ldots + \Theta_q B^q, \quad \Phi_0 = \Theta_0 = I_r, \tag{3}$$

I_r is an $r \times r$ identity matrix, and the diagonal matrix $\text{diag}(\nabla^d)$ is defined as:

$$diag(\nabla^d) = \begin{bmatrix} \nabla^{d_1} & 0 & \ldots & 0 \\ 0 & \nabla^{d_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \nabla^{d_r} \end{bmatrix}, \tag{4}$$

where $\nabla = 1 - B$, and $d_i \in (-1/2, 1/2)$, for all $i = 1, 2, \ldots, r$. Here, $Z_t = (z_{1,t}, \ldots, z_{r,t})^\top$ in eq. (2) is an r-dimensional independent and identically distributed (i.i.d.) white noise process with a nonsingular covariance matrix Σ. The VARFIMA format naturally combines the feature of a univariate ARFIMA model and that of a VARMA process, thus providing a flexible modeling framework for empirical applications. See Sowell (1989) about the structural VARFIMA model.

We theoretically need to evaluate the autocovariance function of W_t before we can compute the conditional log-likelihood function exactly. This task is not trivial for the VARFIMA process. As clearly pointed out in Tsay (2010a), the presence of the AR parameters greatly complicates the computation of the autocovariance functions when Sowell's (1989) algorithm is employed, because it involves hypergeometric functions that need to be evaluated with a truncated infinite sum. A rounding error is inevitable from using Sowell's (1989) methodology. The same problem also applies to the program of Dueker and Startz (1998), because they completely follow the approach of Sowell (1989). Of course, the seriousness of the rounding error problems increases with the dimensionality of the model in eq. (2). What is worse is that even for the model with $r = 2$, Box-Steffensmeier and Tomlinson (2000, p. 71) observe that the program of Dueker and Startz (1998) is extremely sensitive to the starting values and their computation with the Congressional approval data of DGW (1997) tends to get stuck in local minima. These observations call for a new way to estimate the model in eq. (1) so as to enhance the usefulness of this interesting model. In the next section we discuss the rationale of applying the CLDL algorithm to the structural VARFIMA model.

3. CLDL algorithm of Tsay (2010a)

This section explains that the task of estimating the model in eq. (1) is much less daunting if we want to impose the following condition in Tsay (2010a).

Assumption A. $\Phi(B)$ *is diagonal.*

Under Assumption A, W_t in eq. (2) can be rewritten as:

$$\Phi(B)W_t = diag(\nabla^{-d})\Theta(B)Z_t. \quad (5)$$

Since the autocovariance function of $\text{diag}(\nabla^{-d})\Theta(B)Z_t$ is evaluated exactly with the results in eq. (13) of Tsay (2010a), we immediately can evaluate the likelihood function of W_t if $\Phi(B) = I_r$. Following Tsay (2010a), we employ Whittle's (1963) multivariate Durbin–Levinson algorithm to speed up the computation, and the exact conditional likelihood function of W_t given Ψ and $\Phi(B) = I_r$ is:

$$L(W_t|\Psi,\Phi(B) = I_r) = (2\pi)^{\frac{-rT}{2}}\left\{\prod_{j=1}^{T} det(R_{j-1})\right\}^{\frac{-1}{2}} \times exp\left\{-\frac{1}{2}\sum_{j=1}^{T}\left(W_j-\widehat{W}_j\right)^{\top}R_{j-1}^{-1}\left(W_j-\widehat{W}_j\right)\right\}, \quad (6)$$

where $\widehat{W}_j$ denotes the one-step ahead predictor of W_j with the observation $W(j-1) = (W_1^{\top}, W_2^{\top}, \ldots, W_{j-1}^{\top})^{\top}$ as $j \geq 2$. Here, R_{j-1} is the corresponding one-step ahead prediction error matrix. As $j = 1$, $\widehat{W}_1 = 0$, and $R_0 = \Omega(0) = E(W_t W_t^{\top})$. For the definition and computation of $\widehat{W}_j$ and those of R_{j-1}, see Whittle (1963) or Proposition 11.4.1 of Brockwell and Davis (1991), or Section 2 of Tsay (2010a) for the details.

If $\Phi(B)$ is not an identity matrix, then we apply the idea of the conditional likelihood function in Box and Jenkins (1976, Chapter 7) again as the way we deal with the structural parameters Ψ to obtain W_t. In other words, we transform W_t into $\text{diag}(\nabla^{-d})\Theta(B)Z_t$ for a given choice of parameters in $\Phi(B)$ and of suitable starting values. Particularly, if $p = 1$, then conditional on Φ_1 and W_1, $W_t - \Phi_1 W_{t-1}$, $t = 2, 3, \ldots, T$, is a VARFIMA$(0, d, q)$ process, and we denote its associated conditional likelihood function as:

$$L(W_t|\Psi,\Phi_1,W_1) \equiv L(W_t - \Phi_1 W_{t-1}|\Psi), \quad for \quad p = 1, t = 2, 3, \ldots, T. \quad (7)$$

Applying the multivariate Durbin–Levinson algorithm to the transformed data, $W_t - \Phi_1 W_{t-1}$, we simultaneously estimate all the parameters of the structural VARFIMA model with the numerical optimization method. As a consequence, no rounding error occurs during the evaluation of the conditional likelihood function of the structural VARFIMA model, no matter whether $\Phi(B) = I_r$ or not.

It is clear that the procedure of estimating the structural VARFIMA model is to combine Whittle's (1963) multivariate Durbin–Levinson algorithm with the autocovariance function displayed in eq. (13) of Tsay (2010a). It is efficient in computation and is not subject to any rounding error. Furthermore, interested users can follow this procedure and implement it with standard statistics packages.

The limitation of the CLDL algorithm is on the restrictions imposed in Assumption A. Basically, the algorithm is unable to deal with the cases where the off-diagonal elements of the AR matrices are not zero. As mentioned previously, a rounding error is present if we employ Sowell's (1989) algorithm under this circumstance. To the best of our knowledge, there is no algorithm that can evaluate the likelihood function of such a general model exactly. However, Tsay (2010a) argues that the condition in Assumption A is not stringent at all if we look back to the development of the VARFIMA literature. It is well known that the number of parameters of VARFIMA processes increases at the rate of r^2 for an additional value of p or q. Accordingly, for ease of computation and without loss of the parsimonious principle of Box and Jenkins (1976), it is natural to adopt a simplified version of the VARFIMA model when the dimensionality of the data series is large. For example, Haslett and Raftery (1989) assume a homogeneous structure on the fractional differencing and ARMA parameters across meteorological stations to describe the wind speeds recorded at 12 synoptic meteorological stations in Ireland when using the VARFIMA model for their spatial data. In other words, not only do they impose Assumption A, but they also require a much more restrictive MA structure than we do in Assumption A.

From a deeper theoretical point of view, when Assumption A is satisfied, the model in eq. (2) can be further expressed as:

$$\Theta(B)^{-1}diag(\nabla^{d})\Phi(B)W_t = Z_t, \quad (8)$$

i.e., W_t in eq. (8) is a form of VAR(∞) process that is flexible enough to capture the major feature of many multivariate time series. Indeed, we can further rewrite the model in eq. (8) as:

$$\Sigma^{-1/2}\Theta(B)^{-1}diag(\nabla^{d})\Phi(B)W_t = M_t, \quad (9)$$

where M_t is an i.i.d. vector white noise. Consequently, the model in eq. (9) is really general enough to encompass many high dimensional time series models.

4. Congressional approval and economic expectations

This section applies the CLDL algorithm to the model in eqs. (1) and (2) with the data of DGW (1997). First, the dataset has been adopted by Box-Steffensmeier and Tomlinson (2000) to check whether Congressional Approval and Economic Expectations are fractionally cointegrated. This dataset also has been used by Tsay (2010b) to study the divided party government hypothesis. This quarterly data span from 1974:1 to 1993:4. Fig. 1 displays the movements of these two time series. The result of Box-Steffensmeier and Tomlinson (2000) is treated as the benchmark of our empirical studies. Second, the length of this data is only 80, which reflects the observation in Lebo et al. (2000, p. 38) about the small sample size of most political time series. However, we show that various bivariate structural VARFIMA models still can be successfully estimated with such a small sample size by the CLDL algorithm. Third, this dataset contains the observations of presidential approval, and thus we can also test the divided party government hypothesis by adding the data of presidential approval to the above two series in the next section. Doing so not only allows us to check the robustness of the findings generated from the bivariate VARFIMA models, but also demonstrates the power of the CLDL algorithm in estimating a higher dimensional time series model with limited observations.

As mentioned previously, Box-Steffensmeier and Tomlinson (2000) do not report any one-step maximum likelihood estimates. They adopt a two-step procedure, i.e., they first run a cointegration regression and then estimate the order of integratedness of the resulting ordinary least squares (OLS) residuals as 0.4. However, they feel their testing results are unsatisfactory, because a large standard error of 0.45 is found. In other words, they obtain a fractional cointegration between Congressional approval and economic expectations, but this cointegration relationship is also consistent with the usual cointegration model with a short memory error term if we take the large standard error into account. Box-Steffensmeier and Tomlinson (2000, p. 72) explain this large standard error as being expected since there are only 80 observations. On the contrary, the CLDL algorithm is a one-step procedure, and all the short-run and long-run dynamic parameters are estimated simultaneously. Since we do not need to follow the two-step procedure of Box-Steffensmeier and Tomlinson (2000) to estimate the fractional differencing parameter of the residuals from the first-step OLS estimation, the potential small sample bias from the first step estimation can be completely avoided with the use of the CLDL algorithm.

Before presenting our estimation results, we summarize some important findings in Table 1 about the long-run effect of economic expectations on Congressional approval generated from different models, including the autoregressive distributed lag (ADL), the fractional cointegration, the stochastic linear difference equation (SLDE), and the well-known VAR models. The magnitude of the long-run effect should be the focus of the studies as it is related to whether the economic prospects of the public profoundly affect their support of the Congress. The maximum magnitude found in the 9 estimates of Table 1 is 0.49, and most of them lie between 0.24 and 0.36.

Defining Congressional approval and economic expectations at time t as ConAppl_t and EconExp_t, respectively, we adopt the following specification for the empirical applications:

$$\begin{bmatrix} \nabla^{d_1} & 0 \\ 0 & \nabla^{d_2} \end{bmatrix} \begin{bmatrix} 1-\Phi_{11,1}B & 0 \\ 0 & 1-\Phi_{22,1}B \end{bmatrix} \times \begin{bmatrix} ConAppl_t - \beta_1 EconExp_t - \alpha_1 \\ EconExp_t - \gamma_1 EconExp_{t-1} - \alpha_2 \end{bmatrix} = \tilde{W}_t, \tag{10}$$

where

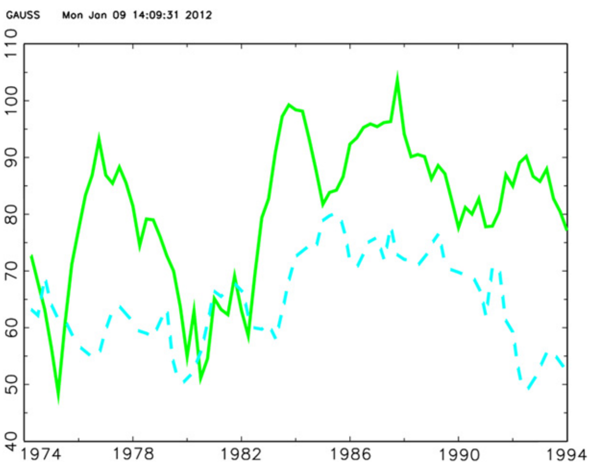

Fig. 1. Congressional approval and economic expectations. Note: Solid line denotes congressional approval, while dashed line represents economic expectations.

Table 1
Summary results for the long-run effects of economic expectations on congressional approval.

Authors	Data	Model	Long-run effects
Box-Steffensmeier and Tomlinson (2000)	Quarterly: 1974–1993	ARFIMA	0.24
De Boef and Keele (2008)	Quarterly: 1974–1993	ADL and and ECM	0.34 from ADL 0.36 from ECM
Chanley et al. (2000)	Quarterly: 1980–1997	VAR	0.25
DGW (1997)	Quarterly: 1974–1993	SLDE	0.35
Lebo (2008)	Monthly: 1995–2005	ARFIMA	0.307 from Clinton presidency
Ramirez (2009)	Quaterly: 1974–2000	SLDE and ECM	0.125 from SLDE 0.36 from ECM
Rudolph (2002)	Quaterly: 1974–1998	ADL and ECM	0.49 (biggest estimate)

Notes: ADL means the autoregressive distributed lag model, SLDE represents the stochastic linear difference equation, and ECM denotes the error-correction model.

$$\tilde{W}_t = \begin{bmatrix} 1+\Theta_{11,1}B & 0 \\ 0 & 1+\Theta_{22,1}B \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix}, \quad Var\begin{bmatrix} u_t \\ v_t \end{bmatrix} = \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}. \tag{11}$$

The optimization algorithm used to implement the CLDL algorithm is the quasi-Newton algorithm of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) contained in the GAUSS MAXLIK library. Furthermore, choosing the starting values for the CLDL algorithm to estimate the following 4 VARFIMA models is pretty easy. Indeed, the corresponding estimation results are not sensitive to the starting values as found by Box-Steffensmeier and Tomlinson (2000) when using the code of Dueker and Startz (1998). The GAUSS program will be posted on the website: http://idv.sinica.edu.tw/wjtsay/htm/jen02.htm

We are particularly interested in two scenarios depending on whether γ_1 in eq. (10) is equal to 1 or not. This distinction is important, because EconExp_t is nonstationary when $\gamma_1 = 1$, and the associated relationship $\text{Approval}_t - \beta_1 x\, \text{EconExp}_t - \alpha_1$ corresponds to the well-known fractional cointegration model. On the other hand, when $\gamma_1 \neq 1$ and $\gamma_1 < 1$, EconExp_t is a stationary fractionally integrated process and its order of integratedness is $d_2 < 1/2$.

When $\gamma_1 = 1$ is imposed, we have:

$$\begin{bmatrix} \nabla^{-0.1056}(1-0.8660B)(ConAppl_t - 0.5718\, EconExp_t - 18.8473) \\ \nabla^{0.0229}(1+0.3290B)(EconExp_t - EconExp_{t-1} + 0.1575) \end{bmatrix} = \tilde{W}_t^{\gamma_1=1}, \tag{12}$$

where

$$\tilde{W}_t^{\gamma_1=1} = \begin{bmatrix} 1+0.2982B & 0 \\ 0 & 1+0.5001B \end{bmatrix} \begin{bmatrix} u_t^{\gamma_1=1} \\ v_t^{\gamma_1=1} \end{bmatrix}, \quad Var\begin{bmatrix} u_t^{\gamma_1=1} \\ v_t^{\gamma_1=1} \end{bmatrix} = \begin{bmatrix} 20.0087 & -15.5807 \\ -15.5807 & 26.2968 \end{bmatrix}. \tag{13}$$

This implies the differencing parameter of EconExp_t is 1.0229 and is consistent with the results in Table 1 of Box-Steffensmeier and Tomlinson (2000) where EconExp_t is likely to be a nonstationary process.

To check the robustness of the above findings, we re-estimate the model in eq. (10), but γ_1 is no longer equal to 1. The CLDL algorithm produces:

$$\begin{bmatrix} \nabla^{0.0145}(1-0.7836B)(ConAppl_t - 1.0642\, EconExp_t + 18.3666) \\ \nabla^{0.1017}(1+0.2829B)(EconExp_t - 0.8745\, EconExp_{t-1} - 9.4381) \end{bmatrix} = \tilde{W}_t^{\gamma_1\neq 1}, \tag{14}$$

where

$$\tilde{W}_t^{\gamma_1\neq 1} = \begin{bmatrix} 1+0.2768B & 0 \\ 0 & 1+0.4571B \end{bmatrix} \begin{bmatrix} u_t^{\gamma_1\neq 1} \\ v_t^{\gamma_1\neq 1} \end{bmatrix}, \quad Var\begin{bmatrix} u_t^{\gamma_1\neq 1} \\ v_t^{\gamma_1\neq 1} \end{bmatrix} = \begin{bmatrix} 40.7109 & -27.3292 \\ -27.3292 & 24.8139 \end{bmatrix}. \tag{15}$$

The inference results for eqs.(12)–(15) are presented in Table 2.

Before discussing the details of the estimation findings, we first illustrate the computing cost of the CLDL algorithms for these two models. It takes 188 iterations for the model in eqs. (12) and (13) to achieve normal convergence, while the number is 172 for the model in eqs. (14) and (15). Nevertheless, the computing time for both models is less than 100 s, even though we estimate 13 parameters for the model in eqs. (14) and (15). As compared to Dueker and Startz (1998) who demonstrate that it takes 35 min for each iteration of their bivariate VARFIMA process with 121 observations and 18 parameters when implementing Sowell's (1989) algorithm, the CLDL algorithm surely shows its superiority in computational efficiency.

The estimation results in Table 2 are essentially qualitatively similar to those in Box-Steffensmeier and Tomlinson (2000) even though EconExp_t is imposed as being stationary. In particular, eq. (14) presents the estimate of γ_1 as 0.8745, which is close to the margin of a unit root specification. When comparing the estimated covariance matrix in eq. (13) with that in eq. (15), we find the nonstationary economic expectations specification seems to be promising in explaining the time series behaviors of the data, because the variation left in (u_t, v_t) is relatively smaller. Given the above estimation results and the findings in Table 1 of Box-Steffensmeier and Tomlinson (2000), we tentatively pay greater attention to the results generated from the fractional cointegration model in eqs. (12) and (13). Nevertheless, we clarify here that, in the literature there exists no well-accepted theorem to address the selection of model specification of the structural VARFIMA models, and further efforts need to be devoted to this line of research. The issue of model specification surely is crucial to the empirical outcomes of this paper, but we feel

Table 2
Estimates of parameters from the structural VARIMA model under stationary or nonstationary economic expectations.

Parameter	Stationary economic expectations			Nonstationary economic expectations		
	Estimate	S.E.	$\lvert t-\text{Ratio}\rvert$	Estimate	S.E.	$\lvert t-\text{Ratio}\rvert$
d_1	0.0145	0.1391	0.1042	−0.1056	0.1106	0.9547
d_2	0.1017	0.1325	0.7675	0.0229	0.1261	0.1816
$\Phi_{11,1}$	0.7836	0.0852	9.1971	0.8660	0.0634	13.6593
$\Phi_{22,1}$	−0.2829	0.2135	1.3250	−0.3290	0.2935	1.1209
$\Theta_{11,1}$	0.2768	0.1169	2.3678	0.2982	0.1082	2.7560
$\Theta_{22,1}$	0.4571	0.2200	2.0777	0.5001	0.2644	1.8914
Σ_{11}	40.7109	38.0619	1.0696	20.0087	9.2378	2.1660
Σ_{22}	24.8139	3.9823	6.2310	26.2968	4.1960	6.2671
Σ_{12}	−27.3292	17.8605	1.5301	−15.5807	8.0597	1.9331
α_1	−18.3666	49.9033	0.3680	18.8473	21.8003	0.8645
α_2	9.4381	5.1003	1.8504	−0.1575	0.7209	0.2184
γ_1	0.8745	0.0678	12.8982	1.000	–	–
β_1	1.0642	0.6785	1.5684	0.5718	0.2897	1.9737

Notes: The results are based on the models in eqs. (10) and (11) under two different specifications on the value of γ_1. S.E. denotes standard errors. The *t*-ratio is computed from using the information matrix based on the numerical Hessian matrix.

this problem should not downsize the contribution of this paper too much, as the objective of this paper is to show the computing advantages of the CLDL algorithm, rather than to argue what kind of structural model best fits the data of DGW (1997).

Table 2 also reveals that the estimates of d_1 and d_2 are close to 0, when $\gamma_1 = 1$ is imposed. The value of $d_2 = 0.0229$ indicates that the differencing parameter of the economic expectations is 1.0229, which is close to the value of 0.86 found in Table 1 of Box-Steffensmeier and Tomlinson (2000). Moreover, the one-step estimate of $d_1 = -0.1056$ is not significant at the 5% level of significance and is qualitatively identical to the finding in Table 2 of Box-Steffensmeier and Tomlinson (2000) where the differencing parameter of the residuals from the first-step OLS estimate is 0.40, but is not significantly different from zero, either. Overall, the evidence of fractional cointegration is not strong with the data of DGW (1997), no matter whether we use the one-step CLDL algorithm, or the two-step procedure of Box-Steffensmeier and Tomlinson (2000). Consequently, the observation of this paper is similar to that in Box-Steffensmeier and Tomlinson (2000), i.e., the relationship between economic expectations and Congressional approval is more consistent with the usual cointegration model with a short memory disturbance term, if we employ a bivariate structural VARFIMA model.

Table 2 also displays that the standard error from estimating d_1 is 0.1106, which is far less than 0.45 found in Table 2 of Box-Steffensmeier and Tomlinson (2000). As compared to the two-step procedure used in Box-Steffensmeier and Tomlinson (2000), the CLDL algorithm provides a narrower confidence interval for testing the parameters of interest. A smaller confidence interval is certainly welcome for empirical users, because it helps us pin down a closer location of the true parameter value. Although we realize further works should be done to address the important model specification issue of VARFIMA models, from the standard knowledge of the time series analysis, the abovementioned smaller confidence interval is expected, because a one-step procedure is usually more efficient than a two-step procedure and is a consistent estimator if the model is correctly specified. As a consequence, the inference results confirm the potential of using the structural VARFIMA model in analyzing the interrelation between Congressional approval and economic expectations.

A notable competitor for estimating the bivariate VARFIMA model is the semiparametric QMLE of Lobato (1999) who develops a two-step estimator based on an extension of the objective function considered in Robinson (1995). However, Lebo et al. (2000, p. 38) point out that the semiparametric methods have undesirable small sample properties, this is unfortunate given the small size of most political time series. For example, the data used in this paper is only 80. Another comment made in footnote 19 of Lebo et al. (2000) is that both Robinson's and Lobato's procedures do not allow for the estimation of short- and long-term dynamics simultaneously. Indeed, the intention of this paper is to introduce a fast and reliable algorithm to estimate the short- and long-term dynamics of a structural VARFIMA process simultaneously. Since Tsay (2010a, p.742) shows that the performance of the CLDL algorithm is much better than that of the semiparametric QMLE of Lobato (1999) given that the model is correctly specified, the introduction of the CLDL algorithm for the structural VARFIMA model helps resolve the concern of Lebo et al. (2000, p. 38) that "the only complaints about full maximum likelihood estimation concern its computationally intensive algorithm".

Another notable competitor for testing the parameters of the structural VARFIMA model is bootstrapping. Andersson and Gredenhoff (1998) consider the issue of testing the parameter of a univariate ARFIMA type model. Nevertheless, one may need to put forth more efforts to extend their paper to the multivariate ARFIMA process, because it still requires a huge computational cost to test a specific value of a fractional differencing parameter of a univariate ARFIMA model with the bootstrap method. Particularly, under the VARFIMA framework, if you want to test the value of the first differencing parameter, d_1, we need to estimate the other $r-1$ differencing parameters in order to obtain the residuals for later resampling purposes. For example, Andersson and Gredenhoff (1998) adopt the well-known GPH method to implement their test. From the argument of Lebo et al. (2000, p. 38) that the semi-parametric methods have undesirable small sample properties, we know this issue is serious given the small size of most political time series. Nevertheless, this small sample bias problem is even more serious when we have to estimate a large number of differencing parameters (for example $r-1$) simultaneously, if we adopt the procedure of Lobato (1999) under this circumstance. The other drawback of the method of Lobato (1999) is that it cannot estimate the short-term dynamics characterized by the AR and MA parameters. Hence, we need a second estimation procedure to apply to the residuals generated from the fractionally differenced series using the fractional differencing parameters ($r-1$ as mentioned previously) estimated from the first step estimation. In short, we need two estimation steps before we can get the residuals to start the bootstrap operation. Therefore, the small sample bias might be more serious for the bootstrap procedure under the VARFIMA framework, let alone the computational burden of the bootstrap certainly increases with the number of replications to construct the bootstrapped confidence interval. To the best of our knowledge, we find no method to address the bootstrap testing for the structural VARFIMA models at all. On the other hand, the one-step CLDL algorithm can produce the t-ratio based on the Hessian matrix calculation of the information matrix as we have done in this paper. It is easy to conduct hypothesis testing with the CLDL algorithm.

To further check the robustness of the findings in eqs. (12) and (13), we re-estimate this model, but $p = 2$ is used instead. The CLDL algorithm produces:

$$\begin{bmatrix} \nabla^{-0.4997}(1-1.5433B+0.5879B^2)(ConAppl_t - 0.5231\,EconExp_t - 22.0893) \\ \nabla^{-0.3632}(1-0.3705B-0.1381B^2)(EconExp_t - EconExp_{t-1} - 0.1390) \end{bmatrix} = \tilde{W}_t^{\gamma_1=1}, \quad (16)$$

where

$$\tilde{W}_t^{\gamma_1=1} = \begin{bmatrix} 1-0.0069B & 0 \\ 0 & 1+0.1545B \end{bmatrix} \begin{bmatrix} u_t^{\gamma_1=1} \\ v_t^{\gamma_1=1} \end{bmatrix}, \quad \begin{bmatrix} u_t^{\gamma_1=1} \\ v_t^{\gamma_1=1} \end{bmatrix} = \begin{bmatrix} 17.7293 & -13.3838 \\ -13.3838 & 25.6861 \end{bmatrix}. \quad (17)$$

Table 3 displays the corresponding inference results.

Table 3 shows that the estimates of d_1 and d_2 are both close to 0 and not significant at the 5% level of significance. This indicates that the evidence of fractional cointegration is not strong with the data of DGW (1997) again, and the findings in Table 2 of Box-Steffensmeier and Tomlinson (2000) and those in Table 2 of this paper remain robust with a more general model specification. We also find that the long-run effect of Congressional approval on economic expectations in eq. (16) is 0.5231, which is very close to the value of 0.5718 found in eq. (12). Accordingly, the use of the VARFIMA$(2,d,1)$ model provides an additional robust checking on the results from the VARFIMA$(1,d,1)$ counterpart.

Table 3
Estimates of parameters under nonstationary economic expectations and VARFIMA(2, d, 1) specification.

Parameter	Nonstationary economic expectations		
	Estimate	S.E.	$\lvert t-\text{Ratio}\rvert$
d_1	−0.4997	0.5802	0.8613
d_2	−0.3632	0.5676	0.6399
$\Phi_{11,1}$	1.5433	0.3695	4.1767
$\Phi_{11,2}$	−0.5879	0.3243	1.8128
$\Phi_{22,1}$	0.3705	0.5432	0.6821
$\Phi_{22,2}$	0.1381	0.2017	0.6847
$\Theta_{11,1}$	−0.0069	0.3319	0.0208
$\Theta_{22,1}$	0.1545	0.4368	0.3537
Σ_{11}	17.7293	8.1401	2.1780
Σ_{22}	25.6861	4.4268	5.8024
Σ_{12}	−13.3838	8.2393	1.6244
α_1	22.0893	23.4969	0.9401
α_2	0.1390	0.5027	0.2765
γ_1	1.000	–	–
β_1	0.5231	0.2962	1.7660

Notes: The results are based on the models in eqs. (16) and (17) where the value of γ_1 is imposed to be 1. S.E. denotes standard errors. The t-ratio is computed from using the information matrix based on the numerical Hessian matrix.

5. Congressional and Presidential approval

This section tests the divided party government hypothesis based on a 3-dimensional structural VARFIMA model using ConAppl$_t$, EconExp$_t$, and the data of Presidential Approval (PreAppl$_t$, hereafter) from DGW (1997). The main objective is to demonstrate the power of the CLDL algorithm in estimating a higher dimensional time series model with a small sample size. In the literature, PreAppl$_t$ has been included as one of the regressors in explaining macropartisanship by Maestas (1998) using the ARFIMA model.

In the political science literature, Lebo (2008) employs a multivariate ARFIMA and establishes a strong and positive relationship between Congressional approval and lagged presidential approval with the monthly data ranging from 1995 to 2005. Nevertheless, the analysis of Lebo (2008) is a two-step procedure, because the original data used in Lebo (2008) need to be differenced using estimated fractional differencing parameters. Thus, the results of Lebo (2008) might be subject to bias induced from the first-stage differencing parameter estimation. On the other hand, our paper adopts a one-step CLDL algorithm that can test the long-run influence of PreAppl$_t$ on ConAppl$_t$ in one step.

The major feature of this section is to treat PreAppl$_t$, ConAppl$_t$, and EconExp$_t$ as endogenous, and we use the structural VARFIMA model to embrace this idea. The theoretical justification is not new. First, as pointed out in

Tsay (2010b, p. 138), by the divided party government hypothesis, competing parties adopt strategies to countermeasure the action of the opposite party if this hypothesis is truly in the mind of political actors. Presidential approval and Congressional approval by nature are the functions of political action. Thus, $ConAppl_t$ and $PreAppl_t$ are endogenous with each other. This implies these two approvals should not have a clear causal relationship as specified in DGW (1997) and De Boef and Keele (2008).

Second, Tsay (2010b) cite the arguments of DGW (1997, p. 186) that "because citizens hold the president accountable for the state of the economy, economic evaluations will affect the president's standing among the public. We believe the same holds for Congress." This is why DGW (1997) use $EconExp_t$ to explain Congressional approval. Moreover, De Boef and Kellstedt (2004, p. 648) argue that consumer confidence should not solely be treated as a right-hand side (RHS) variable in analyzing the political economy. Politics also affects economics.

The above statements jointly indicate that we should treat $PreAppl_t$, $ConAppl_t$, and $EconExp_t$ as endogenous. Moreover, by the arguments in Lebo (2008, p. 8) that the Presidential approval series is nearly a unit-root and clearly has long memory, we thus extend the models in eqs. (10) and (11) as:

$$\Phi(B)diag(\nabla^d)\begin{bmatrix} ConAppl_t-\beta_1 EconExp_t-\beta_2 PreAppl_t-\alpha_1 \\ EconExp_t-EconExp_{t-1}-\alpha_2 \\ PreAppl_t-PreAppl_{t-1}-\alpha_3 \end{bmatrix} = \tilde{W}_t, \tag{18}$$

where

$$\Phi(B)diag(\nabla^d) = \begin{bmatrix} \nabla^{d_1}(1-\Phi_{11,1}B) & 0 & 0 \\ 0 & \nabla^{d_2}(1-\Phi_{22,1}B) & 0 \\ 0 & 0 & \nabla^{d_3}(1-\Phi_{33,1}B) \end{bmatrix}, \tag{19}$$

and

$$\tilde{W}_t = \begin{bmatrix} 1+\Theta_{11,1}B & 0 & 0 \\ 0 & 1+\Theta_{22,1}B & 0 \\ 0 & 0 & 1+\Theta_{33,1}B \end{bmatrix}\begin{bmatrix} u_t \\ v_{1t} \\ v_{2t} \end{bmatrix},$$

$$Var\begin{bmatrix} u_t \\ v_{1t} \\ v_{2t} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33} \end{bmatrix}. \tag{20}$$

With the 80 observations from DGW (1997) and the CLDL algorithm, we estimate a 3-dimensional model of 20 parameters. The results are displayed as follows:

$$\begin{bmatrix} \nabla^{-0.2209}(1-0.9165B)(ConAppl_t-0.6010\,EconExp_t-0.1575\,PreAppl_t-17.3393) \\ \nabla^{-0.0454}(1+0.2880B)(EconExp_t-EconExp_{t-1}+0.1899) \\ \nabla^{-0.4135}(1-0.2800B)(PreAppl_t-PreAppl_{t-1}+0.0527) \end{bmatrix} = \tilde{W}_t, \tag{21}$$

where

$$\tilde{W}_t = \begin{bmatrix} 1+0.3187B & 0 & 0 \\ 0 & 1+0.5326B & 0 \\ 0 & 0 & 1+0.0516B \end{bmatrix}\begin{bmatrix} u_t \\ v_{1t} \\ v_{2t} \end{bmatrix}, \tag{22}$$

and

$$Var\begin{bmatrix} u_t \\ v_{1t} \\ v_{2t} \end{bmatrix} = \begin{bmatrix} 25.4521 & -16.8906 & -17.9220 \\ -16.8906 & 26.3388 & 3.3821 \\ -17.9220 & 3.3821 & 44.3772 \end{bmatrix}. \tag{23}$$

Table 4 illustrates the corresponding inference results.

There are several interesting findings in Table 4. Conditional that the presidential approval and economic expectations are both nonstationary, we find the long-run effect of presidential approval on Congressional approval is 0.1575, which is larger than the value, 0.08, found in De Boef and Keele (2008), but less than that of 0.26 in Tsay (2010b) based on a long memory ADL model. Second, this model at the same time shows that the long-run effect of economic expectations on Congressional approval is 0.6010, which is very close to the aforementioned estimates, 0.5231 and 0.5718, obtained from the 2-dimensional VARFIMA$(2,d,1)$ and VARFIMA$(1,d,1)$ models, respectively. Thus, the robustness of the estimation results from the 2-dimensional VARFIMA models is confirmed with a higher dimensional VARFIMA model.

Third, as compared to the results in Table 2 where the estimate of d_1 is -0.1056, the corresponding estimate in Table 4 is smaller, -0.2209. Nevertheless, the standard error from estimating d_1 in Table 4 is 0.1032, which is very close to the value of 0.1106 in Table 2. This is the reason that d_1 is significantly different from 0 for the 3-dimensional VARFIMA model. However, we cannot find such a significant testing result from Table 2 of this paper and from Table 2 of Box-Steffensmeier and Tomlinson (2000). Consequently, this section establishes the first evidence that the error term of the fractional cointegration model of Congressional approval data of DGW (1997) might be a fractionally integrated process.

Table 4
Estimates of parameters based on 3-dimensional VARFIMA (1, *d*, 1) specification.

Parameter	Estimate	S.E.	$\|t-\text{Ratio}\|$
d_1	−0.2209	0.1032	2.1405
d_2	−0.0454	0.1248	0.3638
d_3	−0.4135	0.2392	1.7287
$\Phi_{11,1}$	0.9165	0.0551	16.6334
$\Phi_{22,1}$	−0.2880	0.2694	1.0690
$\Phi_{33,1}$	0.2800	0.3452	0.8111
$\Theta_{11,1}$	0.3187	0.1016	3.1368
$\Theta_{22,1}$	0.5326	0.2032	2.6211
$\Theta_{33,1}$	0.0516	0.2343	0.2202
Σ_{11}	25.4521	24.4524	1.0409
Σ_{22}	26.3388	4.2147	6.2493
Σ_{33}	44.3772	7.3377	6.0478
Σ_{12}	−16.8906	10.4266	1.6200
Σ_{13}	−17.9220	24.8277	0.7219
Σ_{23}	3.3821	4.4018	0.7683
α_1	17.3393	26.0263	0.6662
α_2	−0.1899	0.5969	0.3181
α_3	−0.0527	0.2578	0.2044
β_1	0.6010	0.3526	1.7045
β_2	0.1575	0.5338	0.2951

Notes: The results are based on the models in eqs.(18)–(20), where the values of γ_1 and γ_2 are both imposed to be 1. S.E. denotes standard errors. The *t*-ratio is computed from using the information matrix based on the numerical Hessian matrix.

6. Conclusion

This paper demonstrates that the maximum likelihood estimation of the structural VARFIMA model can be easily estimated with the CLDL algorithm of Tsay (2010a). We apply the CLDL method to the Congressional approval data of DGW (1997) and find a much larger long-run effect of economic expectations on Congressional approval. This indicates that the influence of the economic prospects of the public on their support for Congress might be more important than we thought. In addition, this paper documents that the standard error from estimating the fractional differencing parameter of the error term is less than one third of the value found in Table 2 of Box-Steffensmeier and Tomlinson (2000). This feature is important in empirical studies, because it indicates that the CLDL algorithm might provide a narrower confidence interval to deliver a more precise inference for the parameters of interest.

We further test the divided party government hypothesis by treating economic expectations, presidential approval, and Congressional approval as endoengous variables in order to demonstrate the computing power of the CLDL algorithm for the high dimensional VARFIMA model. The resulting findings support the argument of Patterson and Caldeira (1990) and Tsay (2010b) that presidential approval does play an important role in affecting Congressional approval. Most importantly, the inference from the 3-dimensional VARFIMA model reveals that the error term of the fractional cointegration model of Congressional approval data of DGW (1997) is highly likely a fractionally integrated process. This result is in sharp contrast with that of Table 2 of Box-Steffensmeier and Tomlinson (2000) and also justifies the use of the VARFIMA model to capture the long memory features of the Congressional approval data.

References

Andersson, M.K., Gredenhoff, M.P. (1998). Robust testing for fractional integration using the bootstrap. Working paper.

Baillie, R.T., Bollerslev, T., 1994. Cointegration, fractional cointegration, and exchange rate dynamics. Journal of Finance XLIX, 737–745.

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis, Forecasting, and Control, second ed. Holden-Day, San Francisco.

Box-Steffensmeier, J.M., Tomlinson, A.R., 2000. Fractional integration methods in political science. Electoral Studies 19, 63–76.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods, second ed. Springer-Verlag, New York.

Chanley, V.A., Rudolph, T.J., Rahn, W.M., 2000. The origins and consequences of public trust in government: a time series analysis. Public Opinion Quarterly 64, 239–256.

Cheung, Y.W., Lai, K.S., 1993. A fractional cointegration analysis of purchasing power parity. Journal of Business and Economic Statistics 11, 103–112.

Clarke, H.D., Lebo, M., 2003. Fractional (co)integration and government party support in Britain. British Journal of Political Science 33, 283–301.

De Boef, S., Keele, L., 2008. Taking time seriously. American Journal of Political Science 52, 184–200.

De Boef, S., Kellstedt, P.M., 2004. The political (and economic) origins of consumer confidence. American Journal of Political Science 48, 633–649.

Dueker, M., Startz, R., 1998. Maximum-likelihood estimation of fractionally cointegration with an application to U.S. and Canadian bond rates. Review of Economics and Statistics 80, 420–426.

Durr, R.H., Gilmour, J.B., Wolbrecht, C., 1997. Explaining congressional approval. American Journal of Political Science 41, 175–207.

Haslett, J., Raftery, A.E., 1989. Space-time modelling with long-memory dependence: assessing Ireland's wind power resource. Applied Statistics 38, 1–50.

Lebo, M.J., 2008. Divided government, united approval: the dynamics of Congressional and Presidential approval. Congress & the Presidency 35, 1–16.

Lebo, M.J., Walker, R.W., Clarke, H.D., 2000. You must remember this: dealing with long memory in political analysis. Electoral Studies 19, 31–48.

Lobato, I.N., 1999. A semiparametric two-step estimator in a multivariate long memory model. Journal of Econometrics 90, 129–153.

Maestas, C., 1998. Sensitivity of GARCH Estimates: Effects of Model Specification on Estimates of Macropartisan Volatility. Working paper. Department of Political Science, University of Colorado, Boulder.

Patterson, S.C., Caldeira, G.A., 1990. Standing up for Congress: variations in public esteem since the 1960s. Legislative Studies Quarterly 15, 25–47.

Ramirez, M.D., 2009. The dynamics and partisan conflict on Congressional approval. American Journal of Political Science 53, 681–694.

Robinson, P.M., 1995. Gaussian semiparametric estimation of long range dependence. Annals of Statistics 23, 1630–1661.

Rudolph, T.J., 2002. The economic source of Congressional approval. Legislative Studies Quarterly 27, 577–600.

Sowell, F., 1989. Maximum Likelihood Estimation of Fractionally Integrated Time Series Models. Working paper. Carnegie-Mellon University.

Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics 53, 165–188.

Tsay, W.J., 2010a. Maximum likelihood estimation of stationary multivariate ARFIMA processes. Journal of Statistical Computation and Simulation 80, 729–745.

Tsay, W.J., 2010b. The long memory autoregressive distributed lag model and its application on congressional approval. Electoral Studies 29, 128–143.

Whittle, P., 1963. On the fitting of multivariate autoregressions and the approximate canonical factorization of a special density matrix. Biometrika 40, 129–134.