Additional material

on

Exact maximum likelihood estimation of partially nonstationary vector ARMA models

José Alberto Mauricio

Revised: June 18, 2005

A1. Stationary representation of a bivariate partially nonstationary VAR(1) model

Consider a stochastic process following the VAR(1) model $\mathbf{Y}_t = \mathbf{\Phi}_1 \mathbf{Y}_{t-1} + \mathbf{A}_t$ with $\mathbf{Y}_t = [Y_{t1}, Y_{t2}]'$ (so that $M = 2$), i.e.,

$$\begin{bmatrix} Y_{t1} \\ Y_{t2} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} Y_{t-1,1} \\ Y_{t-1,2} \end{bmatrix} + \begin{bmatrix} A_{t1} \\ A_{t2} \end{bmatrix}. \tag{A.1}$$

Model (A.1) satisfies the partial nonstationarity assumptions given in Section 2 if and only if the two roots of $|\mathbf{I} - \mathbf{\Phi}_1 x| = 0$ are such that $x_1 = 1$ (one unit root) and $|x_2| > 1$ (so that $D = 1$), i.e., if and only if the two eigenvalues $\mu_i = \frac{1}{x_i}$ $(i = 1, 2)$ of the autoregressive matrix

$$\mathbf{\Phi}_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \tag{A.2}$$

in (A.1) are such that $\mu_1 = 1$ and $|\mu_2| < 1$.

Under the mentioned assumption on (A.2), it can be seen that the matrix

$$\mathbf{\Pi} = \mathbf{I} - \mathbf{\Phi}_1 = \begin{bmatrix} 1 - \phi_{11} & -\phi_{12} \\ -\phi_{21} & 1 - \phi_{22} \end{bmatrix} \tag{A.3}$$

has one eigenvalue $(1 - \mu_1)$ equal to zero and the other one $(1 - \mu_2)$ different from zero. Hence, the VEC representation $\nabla \mathbf{Y}_t = -\mathbf{\Pi Y}_{t-1} + \mathbf{A}_t$ for (A.1) has the property that $\text{rank}[\mathbf{\Pi}] = M - D = 1$, and $\mathbf{\Pi}$ can be written as displayed in equations (4) and (6), i.e.,

$$\mathbf{\Pi} = \mathbf{\Lambda B}' = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} [1, \beta_2] = \begin{bmatrix} \lambda_1 & \lambda_1\beta_2 \\ \lambda_2 & \lambda_2\beta_2 \end{bmatrix}. \tag{A.4}$$

If the general transformation described in Section 3 is applied to the VEC form of (A.1), it follows from (15)-(17) that $\bar{\mathbf{C}}^{-1}\bar{\mathbf{Y}}_t = (\bar{\mathbf{C}}^{-1}\bar{\mathbf{H}} - \bar{\mathbf{\Lambda}})\bar{\mathbf{Y}}_{t-1} + \mathbf{A}_t$, i.e.,

$$\bar{\mathbf{Y}}_t = (\bar{\mathbf{H}} - \bar{\mathbf{C}}\bar{\mathbf{\Lambda}})\bar{\mathbf{Y}}_{t-1} + \bar{\mathbf{C}}\mathbf{A}_t, \tag{A.5}$$

where

$$\bar{\mathbf{Y}}_t = \begin{bmatrix} \nabla Y_{t2} \\ Y_{t1} + \beta_2 Y_{t2} \end{bmatrix}, \; \bar{\mathbf{H}} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \; \bar{\mathbf{C}} = \begin{bmatrix} 0 & 1 \\ 1 & \beta_2 \end{bmatrix}, \; \bar{\mathbf{\Lambda}} = \begin{bmatrix} 0 & \lambda_1 \\ 0 & \lambda_2 \end{bmatrix}.$$

It is now shown that under the partial nonstationarity assumption imposed on (A.2) and represented explicitly in (A.4), the two eigenvalues of

$$\bar{\mathbf{H}} - \bar{\mathbf{C}}\bar{\mathbf{\Lambda}} = \begin{bmatrix} 0 & -\lambda_2 \\ 0 & 1 - \lambda_1 - \lambda_2\beta_2 \end{bmatrix} \tag{A.6}$$

are less than one in absolute value, which implies that (A.5) is a stationary VAR(1) model for $\{\bar{\mathbf{Y}}_t\}$.

First, note that the two eigenvalues μ_1 and μ_2 of (A.2) satisfy the following two equations:

$$\text{Trace}[\mathbf{\Phi}_1] = \mu_1 + \mu_2 = \phi_{11} + \phi_{22}. \tag{A.7}$$

$$|\,\mathbf{\Phi}_1\,| = \mu_1\mu_2 = \phi_{11}\phi_{22} - \phi_{12}\phi_{21}. \tag{A.8}$$

Hence, $\mu_1 = 1$ implies from (A.7)-(A.8) that

$$\mu_2 = \phi_{11} + \phi_{22} - 1 = \phi_{11}\phi_{22} - \phi_{12}\phi_{21}, \tag{A.9}$$

showing that the four elements of $\mathbf{\Phi}_1$ in (A.2) must satisfy certain specific restrictions for partial nonstationarity to hold. In particular, $|\,\mu_2\,| < 1$ and the first part of (A.9) imply that

$$|\,\phi_{11} + \phi_{22} - 1\,| < 1\,. \tag{A.10}$$

Second, note that (A.3)-(A.4) imply that $\text{trace}[\mathbf{\Pi}] = (1 - \phi_{11}) + (1 - \phi_{22}) = \lambda_1 + \lambda_2\beta_2$, i.e., $\phi_{11} + \phi_{22} - 1 = 1 - \lambda_1 - \lambda_2\beta_2$. It follows from this equation and equation (A.10) that

$$|\,1 - \lambda_1 - \lambda_2\beta_2\,| < 1\,. \tag{A.11}$$

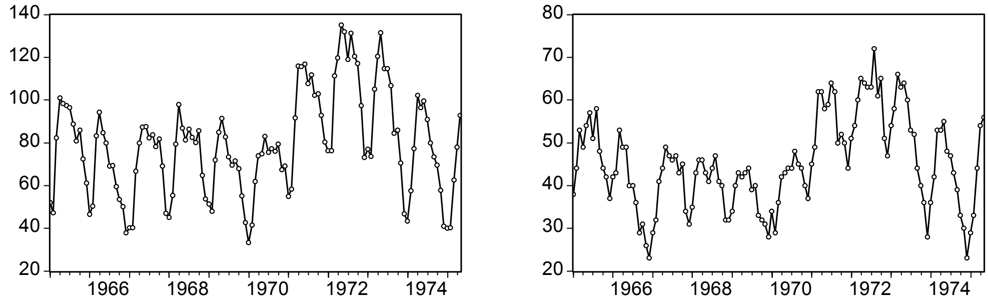
Finally, note that the two eigenvalues of the autoregressive matrix (A.6) are zero and $1 - \lambda_1 - \lambda_2\beta_2$. Hence, (A.11) implies the claimed result that (A.5) is a stationary VAR(1) representation for the partially nonstationary model (A.1), which illustrates the general result following equations (15)-(17) of the article.

A2. The Census Housing Data

Building adequate time series models for seasonally unadjusted monthly data presents some peculiarities that conditional estimation methods often do not handle satisfactorily. This example illustrates the following important points: (i) As in the case of stationary models, EML estimation is clearly preferable to CML in the case of cointegrated systems with possibly noninvertible moving average terms (especially for seasonal models), and (ii) joint and reliable EML estimation of both cointegrating terms and such moving average terms is possible by using the methods described in Section 3 of the article.

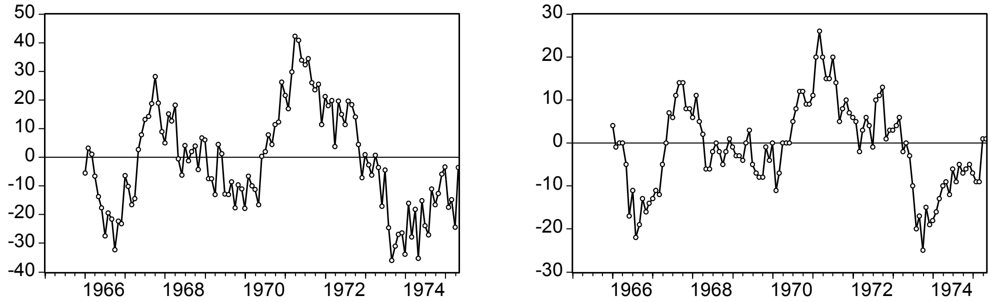
The data used in this example consist of monthly U.S. single-family housing starts (x_{t1}) and houses sold (x_{t2}) over the period January 1965 through May 1975. These data have been considered recently by Mélard et al. (2004, Example 4), Reinsel (1997, Examples 6.4 and 6.6), and Tiao (2001) (who refers to them as the "Census Housing Data"). Point (i) above is illustrated somehow in the last two works, although the possibility mentioned in point (ii) above is considered in none of the three at all. Additionally, the present example illustrates the possibility of estimating common trends in cointegrated systems as a byproduct of EML estimation of a suitable VEC model.

The original (seasonally unadjusted) data shown in Figure A1 exhibit strong seasonal behavior, and so the seasonal differences $y_{t1} = (1 - L^{12})x_{t1}$ and $y_{t2} = (1 - L^{12})x_{t2}$ are considered for further modeling. The seasonally differenced data are displayed in Figure A2, from which it can be seen that $\mathbf{y}_t = [y_{t1}, y_{t2}]'$ is a nonstationary series such that $\nabla \mathbf{y}_t = [\nabla y_{t1}, \nabla y_{t2}]'$ is stationary. Formal unit root tests can be shown to confirm clearly this conclusion obtained by simple visual inspection of the data.



(a) Original Housing Starts Series (b) Original Houses Sold Series

Figure A1. The Original (Seasonally Unadjusted) Census Housing Data (in thousands) over the Period January 1965 Through May 1975.



(a) Seasonally Differenced Housing Starts Series (b) Seasonally Differenced Houses Sold Series

Figure A2. The Seasonally Differenced Census Housing Data (in thousands) over the Period January 1966 Through May 1975.

Hence, following Reinsel (1997, Example 6.6) and Tiao (2001), a model of the form

$$(\mathbf{I} - \mathbf{\Phi}_1 L)\mathbf{Y}_t = (\mathbf{I} - \mathbf{\Theta}_1 L^{12})\mathbf{A}_t \tag{A.12}$$

has been estimated with the bivariate time series $\mathbf{y}_t = [y_{t1}, y_{t2}]'$ shown in Figure A2, and the possibility that the 2×2 matrix $\mathbf{\Pi} = \mathbf{I} - \mathbf{\Phi}_1$ in the

Table A1. Estimation Results for Model (A.12): Seasonally Differenced Census Housing Data. (Estimated standard errors in parentheses.)

	Exact Maximum Likelihood	Conditional Maximum Likelihood
$\hat{\mathbf{\Phi}}_1$	$\begin{bmatrix} 0.4753 & 0.9306 \\ (0.0730) & (0.1359) \\ 0.0974 & 0.7648 \\ (0.0463) & (0.0863) \end{bmatrix}$	$\begin{bmatrix} 0.5033 & 0.8408 \\ (0.0743) & (0.1341) \\ 0.1251 & 0.7109 \\ (0.0463) & (0.0837) \end{bmatrix}$
$\hat{\mathbf{\Theta}}_1$	$\begin{bmatrix} 0.9641 & 0.0000 \\ (0.1480) & (\text{——}) \\ 0.0000 & 1.0844 \\ (\text{——}) & (0.1870) \end{bmatrix}$	$\begin{bmatrix} 0.7309 & 0.0000 \\ (0.0741) & (\text{——}) \\ 0.0000 & 0.7012 \\ (\text{——}) & (0.0748) \end{bmatrix}$
$\hat{\mathbf{\Sigma}}$	$\begin{bmatrix} 29.2020 & \\ 5.5274 & 10.3041 \end{bmatrix}$	$\begin{bmatrix} 38.1903 & \\ 6.7457 & 15.2644 \end{bmatrix}$
Eigenvalues of $\hat{\mathbf{\Pi}} = \mathbf{I} - \hat{\mathbf{\Phi}}_1$	$0.0460, 0.7140$	$0.0523, 0.7335$
Log-likelihood	-663.4662	-669.9050
AIC, BIC	$12.0083, 12.2268$	$12.1233, 12.3418$

corresponding VEC model

$$\nabla \mathbf{Y}_t = -\mathbf{\Pi}\mathbf{Y}_{t-1} + (\mathbf{I} - \mathbf{\Theta}_1 L^{12})\mathbf{A}_t \tag{A.13}$$

be of reduced rank (either $P = 0$ or $P = 1$) has been considered.

Table A1 summarizes the estimation results for model (A.12) obtained through EML and CML (where a few parameters have been set to zero because they were clearly insignificant in a previous estimation run). These results suggest the possibility that $\mathbf{\Pi}$ in (A.13) has a single zero eigenvalue (i.e., that $|\,\mathbf{I} - \mathbf{\Phi}_1 x\,| = 0$ has a single unit root), implying that the two series shown in Figure A2 are cointegrated. To explore this possibility formally, tests based on likelihood ratio statistics have been considered. The tests for the various hypotheses are displayed in Table A2.

Table A2. Tests on the Rank P of $\mathbf{\Pi}$ in Model (A.13) (i.e., on the Number D of Unit Roots of $|\mathbf{I} - \mathbf{\Phi}_1 x| = 0$ in Model (A.12)) Based on Likelihood Ratio Test Statistics.

Hypotheses	Likelihood Ratio Test Statistic	Asymptotic p-value
	Exact Maximum Likelihood	
H_0: $P = 0$ $(D = 2)$ H_1 : $P = 1$ $(D = 1)$	$2 \times [L^*_{\mathrm{E}}(1) - L^*_{\mathrm{E}}(0)] = 59.2195$	Less than 0.01 %
H_0: $P = 1$ $(D = 1)$ H_1 : $P = 2$ $(D = 0)$	$2 \times [L^*_{\mathrm{E}}(2) - L^*_{\mathrm{E}}(1)] = 1.5534$	24.95 %
	Conditional Maximum Likelihood	
H_0: $P = 0$ $(D = 2)$ H_1 : $P = 1$ $(D = 1)$	$2 \times [L^*_{\mathrm{C}}(1) - L^*_{\mathrm{C}}(0)] = 60.0643$	Less than 0.01 %
H_0: $P = 1$ $(D = 1)$ H_1 : $P = 2$ $(D = 0)$	$2 \times [L^*_{\mathrm{C}}(2) - L^*_{\mathrm{C}}(1)] = 1.9866$	18.70 %

NOTES: $L^*_{\mathrm{E}}(P)$ represents the exact log-likelihood computed at the EML estimates of model (A.13) for the three different possible values of $P = \mathrm{rank}(\mathbf{\Pi})$ $(P = 0, 1, 2)$. $L^*_{\mathrm{C}}(P)$ represents the conditional log-likelihood computed at the CML estimates of model (A.13) for the three different possible values of P.

The results in Table A2 indicate that the hypothesis of $P = 0$ (or $D = 2$ unit roots) is strongly rejected, and that the hypothesis of $P = 1$ (or $D = 1$ unit root) can not be rejected. Thus, the likelihood ratio test procedure leads to the conclusion that there is one cointegrating vector and one unit root (i.e., one common stochastic trend) in the bivariate model considered.

Note that both EML and CML lead to the same conclusion, and that this is also the conclusion obtained by Reinsel (1997, Examples 6.4 and 6.6) and by Tiao (2001), who also describe the implications of the possible unit root structure in the seasonal moving average operator (i.e., the implications of the possibility that $\mathbf{\Theta}_1 = \mathbf{I}$), that is clearly present when EML estimation is employed (see Table A1) as opposed to the results obtained through CML.

After the results in Table A2, it seems natural to estimate the VEC model (A.13) under the restriction that $P = \text{rank}(\mathbf{\Pi}) = 1$ (i.e., that $\mathbf{\Pi} = \mathbf{\Lambda B}'$), in which case (see Section 2) (A.13) can be conveniently rewritten as

$$\nabla \mathbf{Y}_t = -\mathbf{\Lambda B}' \mathbf{Y}_{t-1} + (\mathbf{I} - \mathbf{\Theta}_1 L^{12}) \mathbf{A}_t \,, \tag{A.14}$$

where $\mathbf{\Lambda} = [\lambda_1 , \lambda_2]'$ is a vector of adjustment factors, and $\mathbf{B} = [1, \beta_2]'$ is the (normalized) cointegrating vector. Table A3 summarizes the estimation results for model (A.14) obtained through both EML and CML (where the same insignificant parameters than those in Table A1 have been set to zero); see also Figure A3 for a brief diagnostic of the estimated model.

Table A3. Estimation Results for Model (A.14): Seasonally Differenced Census Housing Data. (Estimated standard errors in parentheses.)

	Exact Maximum Likelihood	Conditional Maximum Likelihood
$\hat{\mathbf{\Lambda}}$	$\begin{bmatrix} 0.5191 \\ (0.0710) \\ -0.1085 \\ (0.0461) \end{bmatrix}$	$\begin{bmatrix} 0.4922 \\ (0.0723) \\ -0.1343 \\ (0.0457) \end{bmatrix}$
$\hat{\mathbf{B}}$	$\begin{bmatrix} 1^* \\ -1.8625 \\ (0.0939) \end{bmatrix}$	$\begin{bmatrix} 1^* \\ -1.8223 \\ (0.0975) \end{bmatrix}$
$\hat{\mathbf{\Theta}}_1$	$\begin{bmatrix} 0.9600 & 0.0000 \\ (0.1296) & (\text{——}) \\ 0.0000 & 1.1318 \\ (\text{——}) & (0.1735) \end{bmatrix}$	$\begin{bmatrix} 0.7531 & 0.0000 \\ (0.0663) & (\text{——}) \\ 0.0000 & 0.6918 \\ (\text{——}) & (0.0675) \end{bmatrix}$
$\hat{\mathbf{\Sigma}}$	$\begin{bmatrix} 29.4526 & \\ 5.7112 & 9.9830 \end{bmatrix}$	$\begin{bmatrix} 37.4868 & \\ 7.0347 & 15.9136 \end{bmatrix}$
Eigenvalues of $\hat{\mathbf{\Pi}} = \hat{\mathbf{\Lambda}}\hat{\mathbf{B}}'$	0^* , 0.7212	0^* , 0.7369
Log-Likelihood	−664.2429	−670.8983
AIC, BIC	12.0043, 12.1985	12.1232, 12.3174

NOTE: An asterisk indicates a normalized or an implied parameter value.

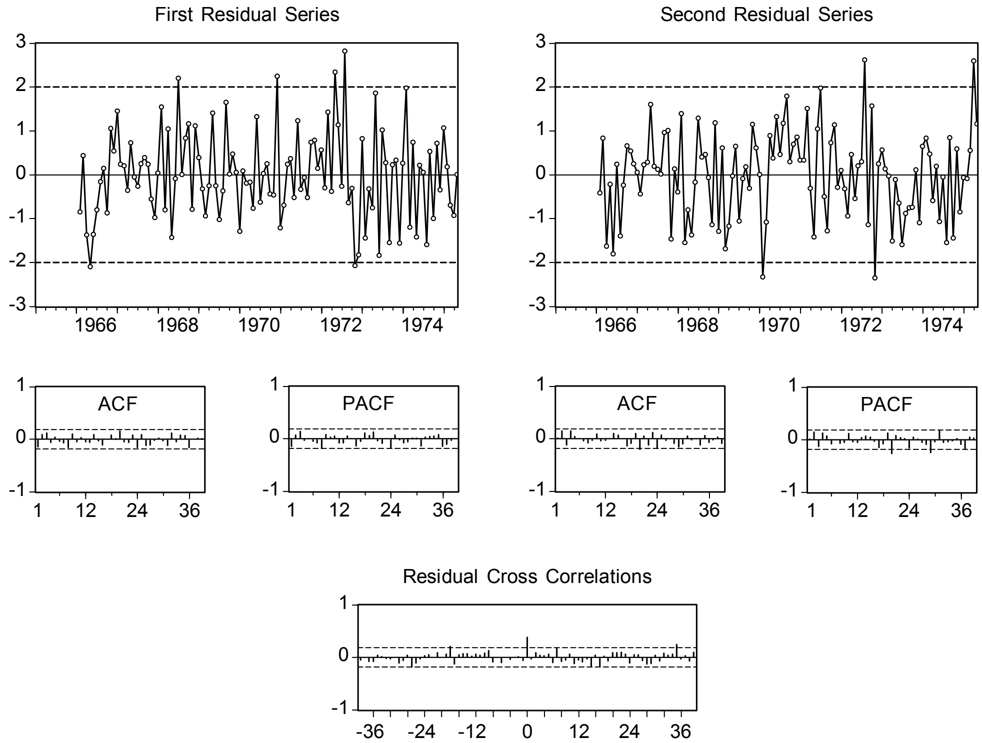


Figure A3. Unconditional Residuals from Exact Maximum Likelihood Estimation of Model (A.14) (See Table A3). (Note: Residual plots are standardized.) When residual simple (ACF) and partial (PACF) autocorrelations, as well as residual cross correlations, are compared to the limits of $\pm 2N^{-0.5} \approx \pm 0.1890$ (with $N = 112$ effective observations), there is no indication of misspecification in the estimated model.

The results given in Table A3 indicate that although the estimated partially nonstationary structure of the VEC model (A.14) is virtually the same irrespective of the estimation method employed, the estimated seasonal moving average and error covariance structures are quite different. In particular, EML estimation implies (i) a possible deterministic seasonal pattern in the original (seasonally unadjusted) census housing data, which is not revealed through CML, as well as (ii) a substantial decrease in the estimated variances of the error processes with respect to CML.

In summary, this example clearly illustrates that, as it often happens in the case of stationary VARMA models, EML estimation of partially nonstationary models can reveal certain important dynamic structure for the original data considered that can not be seen when CML is used instead.

For the sake of completeness, additional issues illustrating EML estimation of common trends are considered now.

In Figure A4(a) the linear combination serie $\hat{w}_t = \hat{\mathbf{B}}'\mathbf{y}_t = y_{t1} - 1.8625 y_{t2}$ is displayed, representing stationary or transitory deviations of $\mathbf{y}_t = [y_{t1}, y_{t2}]'$ from the estimated cointegrating or long-term equilibrium relation $y_{t1} \approx 1.8625 y_{t2}$. Note from Table A3 and equation (A.14) that transitory disequilibria imply significant adjustments in the seasonal differences of both housing starts (negative) and houses sold (positive), that drive the two seasonally differenced series back to their cointegrating relation.

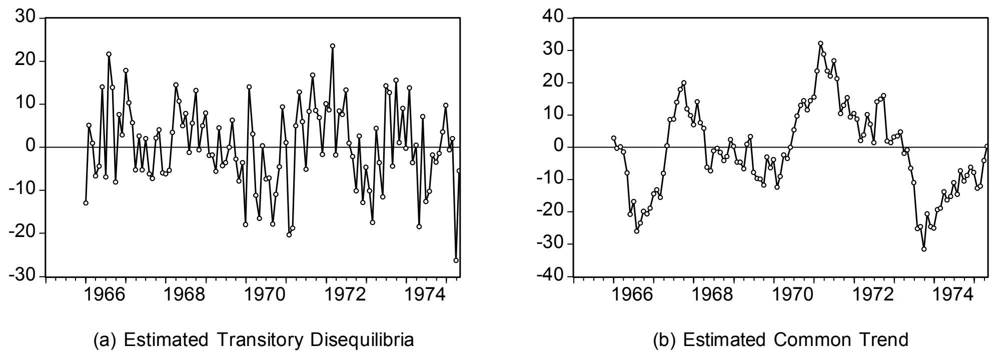


Figure A4. Estimated Transitory Disequilibria and Common Trend for the Seasonally Differenced Census Housing Data (in Thousands) over the Period January 1966 Through May 1975.

Conversely, Figure A4(b) displays the series $\hat{v}_t = \hat{\mathbf{P}}'\mathbf{y}_t = 0.2090 y_{t1} + y_{t2}$, where the normalized 2×1 matrix $\hat{\mathbf{P}} = [0.2090, 1.0]'$ satisfies $\hat{\mathbf{P}}'\hat{\mathbf{\Lambda}} = 0$ (see the end of Section 2), representing the estimated purely nonstationary common trend which is shared by the two series displayed in Figure A2. (Note that the two series $\hat{w}_t$ and $\hat{v}_t$ have been generated using the EML estimates of $\mathbf{\Lambda}$ and $\mathbf{B}$ appearing in Table A3, which are similar to the corresponding CML estimates.)

Hence, defining the 2×2 matrix

$$\hat{\mathbf{Q}} = \begin{bmatrix} \hat{\mathbf{B}}' \\ \hat{\mathbf{P}}' \end{bmatrix} = \begin{bmatrix} 1.0 & -1.8625 \\ 0.2090 & 1.0 \end{bmatrix},$$

it follows that $[\hat{w}_t, \hat{v}_t]' = \hat{\mathbf{Q}}\mathbf{y}_t$ implies that $\mathbf{y}_t = \hat{\mathbf{Q}}^{-1}[\hat{w}_t, \hat{v}_t]'$, so that the two rows of the 2×2 matrix

$$\hat{\mathbf{Q}}^{-1} = \begin{bmatrix} 0.7198 & 1.3406 \\ -0.1504 & 0.7198 \end{bmatrix}$$

provide two sets of coefficients that allow for representing the two series displayed in Figure A2 as linear combinations of the estimated disequilibrium (stationary) series $\hat{w}_t$ and the estimated common trend (purely nonstationary) series $\hat{v}_t$ shown in Figure A4.

In summary, the above operations illustrate the following two general points: (i) it is possible to estimate through EML common trends in cointegrated systems as a byproduct of EML estimation of a suitable VEC model, and (ii) the nonuniqueness of the transformation given by Mélard et al. (2004, Section 3) may pose unnecessary difficulties for the EML estimation process, which, in fact, can be easily overcome through simple normalization after EML estimation whenever estimated common trends are required.