Temporal Disaggregation: Methods, Information Loss, and Diagnostics

Supplementary Materials

Duk B. Jun

KAIST, College of Business, Seoul, Republic of Korea

(*dbjun@business.kaist.ac.kr*)

Jihwan Moon

Warrington College of Business, University of Florida, Gainesville, FL 32611

(*mjhxnna@ufl.edu*)

Sungho Park

W.P. Carey School of Business, Arizona State University, Tempe, AZ 85287

(spark104@asu.edu)

APPENDIX A: OBSERVABILITY CONDITION

A discrete-time system,

$$\begin{aligned} \mathbf{y}_t &= \mathbf{H}_{m\times n}\mathbf{x}_t \\ \mathbf{x}_t &= \mathbf{F}_{n\times n}\mathbf{x}_{t-1} \end{aligned} \tag{1}$$

is completely observable if and only if the observability matrix,

$\mathbf{\Omega} = \begin{pmatrix} \mathbf{H} \\ \mathbf{HF} \\ \mathbf{HF}^2 \\ \vdots \\ \mathbf{HF}^{n-1} \end{pmatrix}$ has rank n (Luenberger, 1979).

Observability is related to the ability to infer what the model is doing in terms of the unique estimation of state variables from a given sequence of observing time series (see Joo and Jun 1997; Jun et al. 2012). The observability matrix must be nonsingular to prevent spurious decomposition. We aim to prove that c must be no more than p+1 to satisfy the observability condition in $\text{ARC}(p,c)$.

A.1 ARC(1,2)

$$\varepsilon_t^c = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{2t-1} \end{pmatrix}.$$

$$\begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{2t-1} \end{pmatrix} = \begin{pmatrix} \phi^2 & 0 \\ \phi & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_{2t-2} \\ \varepsilon_{2t-3} \end{pmatrix} + \begin{pmatrix} 1 & \phi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_{2t} \\ \eta_{2t-1} \end{pmatrix}.$$

$$\mathbf{\Omega} = \begin{pmatrix} 1 & 1 \\ \phi(\phi+1) & 0 \end{pmatrix}.$$

Because $|\mathbf{\Omega}| \neq 0$, the observability condition is satisfied.

A.2 ARC(p,2)

For $p \geq 2$,

$$\varepsilon_t^c = (1 \quad 1 \quad \overbrace{0 \quad \cdots \quad 0}^{p-2})\begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{2t-1} \\ \vdots \\ \varepsilon_{2t-p+1} \end{pmatrix} = \mathbf{H}\begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{2t-1} \\ \vdots \\ \varepsilon_{2t-p+1} \end{pmatrix}, \text{ where } \mathbf{H} = (1 \quad 1 \quad \overbrace{0 \quad \cdots \quad 0}^{p-2}).$$

$$\begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{2t-1} \\ \vdots \\ \varepsilon_{2t-p+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}\begin{pmatrix} \varepsilon_{2t-1} \\ \varepsilon_{2t-2} \\ \vdots \\ \varepsilon_{2t-p} \end{pmatrix} + \begin{pmatrix} \eta_{2t} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F}^2\begin{pmatrix} \varepsilon_{2t-2} \\ \varepsilon_{2t-3} \\ \vdots \\ \varepsilon_{2t-p-1} \end{pmatrix} + \begin{pmatrix} \mathbf{A}_{2\times 2} & \mathbf{O}_{2\times(p-2)} \\ \mathbf{O}_{(p-2)\times 2} & \mathbf{O}_{(p-2)\times(p-2)} \end{pmatrix}\begin{pmatrix} \eta_{2t} \\ \eta_{2t-1} \\ \vdots \\ 0 \end{pmatrix},$$

where $\mathbf{F} = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} 1 & \phi_1 \\ 0 & 1 \end{pmatrix}$, and $\mathbf{O}$ is a zero matrix. $\mathbf{\Omega} = \begin{pmatrix} \mathbf{H} \\ \mathbf{HF}^2 \\ \vdots \\ \mathbf{HF}^{2p-2} \end{pmatrix}$.

To prove that $|\mathbf{\Omega}| \neq 0$, we consider the equation $c_1\mathbf{H} + c_2\mathbf{HF}^2 + \cdots + c_p\mathbf{HF}^{2p-2} = \mathbf{O}$. To show that the rows of $\mathbf{\Omega}$ are linearly independent, we aim to verify that $c_1 = c_2 = \cdots = c_p = 0$ is the only solution of the equation for all $(\phi_1, \phi_2, \cdots, \phi_p)$. Suppose that there is a nonzero solution for the equation. From the equation, $g_1(c_1, c_2, \cdots, c_p)\mathbf{v}_1 + g_2(c_1, c_2, \cdots, c_p)\mathbf{v}_2 + \cdots + g_p(c_1, c_2, \cdots, c_p)\mathbf{v}_p =$ $\mathbf{O}_{1\times p}$ for some $g_1, g_2, ..., g_p$, where $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ are the row vectors of $\mathbf{F}$. Because the determinant of $\mathbf{F}$ is not zero, the rows of $\mathbf{F}$ are linearly independent. Thus, given $(\phi_1, \phi_2, \cdots, \phi_p)$,

$$g_i(c_1, c_2, \cdots, c_p \mid \phi_1, \phi_2, \cdots, \phi_p) = 0 \text{ for } i = 1, \ldots, p.$$

The equations can be transformed into the following system of linear equations:

$$\begin{pmatrix} f_{11}(\phi_1, \cdots, \phi_p) & f_{12}(\phi_1, \cdots, \phi_p) & \cdots & f_{1p}(\phi_1, \cdots, \phi_p) \\ f_{21}(\phi_1, \cdots, \phi_p) & \ddots & & \\ \vdots & & & \vdots \\ f_{p1}(\phi_1, \cdots, \phi_p) & & \cdots & f_{pp}(\phi_1, \cdots, \phi_p) \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ or simply, } \mathbf{RC} = \mathbf{O}.$$

If the rank of **R** is zero, $\phi_i = 0$ for all *i*. This is a contradiction because $\phi_p \neq 0$; therefore, $1 \leq rank(\mathbf{R}) < p$. Then, the dimension of the null space of **R** is $p - rank(\mathbf{R})$, which indicates that $p - rank(\mathbf{R})$ equations of $\phi_1, \phi_2, \cdots, \phi_p$ must be satisfied; therefore, the dimension of the space of $(\phi_1, \phi_2, \cdots, \phi_p)$ is less than *p* for the system to have a nonzero solution. This is a contradiction. Consequently, a nonzero solution does not exist in general.

A.3 ARC(*p,c*)

Case I $p+1 < c$

$$\varepsilon_t^c = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{ct} \\ \vdots \\ \varepsilon_{ct-c+1} \end{pmatrix}.$$

$$\begin{pmatrix} \varepsilon_{ct} \\ \vdots \\ \varepsilon_{ct-c+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \cdots & \phi_p & \mathbf{O}_{1\times(c-p)} \\ 1 & \cdots & 0 & \mathbf{O}_{1\times(c-p)} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & \mathbf{O}_{1\times(c-p)} \end{pmatrix} \begin{pmatrix} \varepsilon_{ct-1} \\ \vdots \\ \varepsilon_{ct-c} \end{pmatrix} + \begin{pmatrix} \eta_{ct} \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F} \begin{pmatrix} \varepsilon_{ct-1} \\ \vdots \\ \varepsilon_{ct-c} \end{pmatrix} + \begin{pmatrix} \eta_{ct} \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F}^c \begin{pmatrix} \varepsilon_{ct-c} \\ \vdots \\ \varepsilon_{ct-2c+1} \end{pmatrix} + \mathbf{A} \begin{pmatrix} \eta_{ct} \\ \vdots \\ \eta_{ct-c+1} \end{pmatrix}$$

for some $c \times c$ matrix **A**.

$$\mathbf{\Omega} = \begin{pmatrix} 1 & \cdots & 1 \\ \mathbf{B} & & \mathbf{O}_{(c-1)\times(c-p)} \end{pmatrix} \text{ for some } (c-1)\times p \text{ matrix } \mathbf{B}.$$

Thus, $\mathbf{\Omega}$ is singular.

Case II $p+1 = c$

$$\varepsilon_t^c = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{ct} \\ \vdots \\ \varepsilon_{ct-c+1} \end{pmatrix}.$$

$$\begin{pmatrix} \varepsilon_{ct} \\ \vdots \\ \varepsilon_{ct-c+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \cdots & \phi_p & 0 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_{ct-1} \\ \vdots \\ \varepsilon_{ct-c} \end{pmatrix} + \begin{pmatrix} \eta_{ct} \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F} \begin{pmatrix} \varepsilon_{ct-1} \\ \vdots \\ \varepsilon_{ct-c} \end{pmatrix} + \begin{pmatrix} \eta_{ct} \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F}^c \begin{pmatrix} \varepsilon_{ct-c} \\ \vdots \\ \varepsilon_{ct-2c+1} \end{pmatrix} + \mathbf{A} \begin{pmatrix} \eta_{ct} \\ \vdots \\ \eta_{ct-c+1} \end{pmatrix}$$

for some $c \times c$ matrix $\mathbf{A}$.

$\mathbf{\Omega} = \begin{pmatrix} 1 & \cdots & 1 \\ \mathbf{B} & \mathbf{O}_{(c-1)\times 1} \end{pmatrix}$ for some $(c-1)\times(c-1)$ matrix $\mathbf{B}$.

With the proof in Appendix A.2, we can verify that $\mathbf{\Omega}$ is nonsingular.

Case III $p+1 > c$

$$\varepsilon_t^c = (\overbrace{1 \quad \cdots \quad 1}^{c}\mathbf{O}_{1\times(p-c)})\begin{pmatrix} \varepsilon_{ct} \\ \vdots \\ \varepsilon_{ct-p+1} \end{pmatrix}.$$

$$\begin{pmatrix} \varepsilon_{ct} \\ \vdots \\ \varepsilon_{ct-p+1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \cdots & \cdots & \phi_p \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}\begin{pmatrix} \varepsilon_{ct-1} \\ \vdots \\ \varepsilon_{ct-p} \end{pmatrix} + \begin{pmatrix} \eta_{ct} \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F}\begin{pmatrix} \varepsilon_{ct-1} \\ \vdots \\ \varepsilon_{ct-p} \end{pmatrix} + \begin{pmatrix} \eta_{ct} \\ \vdots \\ 0 \end{pmatrix} = \mathbf{F}^c\begin{pmatrix} \varepsilon_{ct-c} \\ \vdots \\ \varepsilon_{ct-c-p+1} \end{pmatrix} + \mathbf{A}\begin{pmatrix} \eta_{ct} \\ \vdots \\ \eta_{ct-c+1} \\ \mathbf{O}_{(p-c)\times 1} \end{pmatrix}$$

for some $p \times p$ matrix $\mathbf{A}$.

$|\mathbf{\Omega}| \neq 0$; the proof is the same as in Appendix A.2.

APPENDIX B: PROOFS

In ARC(p,c),

$$\begin{aligned} \phi(L)\varepsilon_t &= \eta_t && \text{for} \quad t = 1,\ldots,cT, \\ \varepsilon_t^a &= \sum_{j=1}^{c}\varepsilon_{t-j+1} && \text{for} \quad t = 1,\ldots,cT, \\ \varepsilon_t^c &= \varepsilon_{ct}^a && \text{for} \quad t = 1,\ldots,T, \end{aligned} \tag{2}$$

where $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\varepsilon_t = 0$ for $t \leq 0$, and $\eta_t \sim NID(0,\sigma_\eta^2)$. L is the lag operator. We verify the following lemmas and Theorem 1 so that the identification procedure can be established for the two-stage process of disaggregation. In the proofs, we carefully treat ε_t^c and ε_{ct}^a because the difference is subtle between the two notations with regard to the lag operator:

$$\nabla\varepsilon_{ct}^a = (1-L)\varepsilon_{ct}^a = \varepsilon_{ct}^a - \varepsilon_{ct-1}^a, \text{ and } \nabla\varepsilon_t^c = (1-L)\varepsilon_t^c = \varepsilon_t^c - \varepsilon_{t-1}^c = \varepsilon_{ct}^a - \varepsilon_{ct-c}^a.$$

Lemma 1. ε_t ~ I(d) if and only if ε_t^a ~ I(d).

Proof. The proof is straightforward from system (2).

Lemma 2. Suppose that ε_t^a and ε_t^c are integrated processes. Then, ε_t^a is stationary if and only if ε_t^c is stationary.

Proof. Suppose that ε_t^a is a zero mean stationary process. By Wold's decomposition, $\varepsilon_t^a =$ $\psi(L)\xi_t$, where $\psi(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \cdots$, $\sum_{j=0}^{\infty}\psi_j^2 < \infty$ with $\psi_0 = 1$, and ξ_t is white noise. We have $\varepsilon_{ct}^a = \psi(L)\xi_{ct}$; thus, $E[\varepsilon_t^c] = E[\psi(L)\xi_{ct}] = 0$ and $\text{Var}[\varepsilon_t^c] = \text{Var}[\psi(L)\xi_{ct}] = \sum_{j=0}^{\infty}\psi_j^2 \text{Var}[\xi_{ct}] = \sum_{j=0}^{\infty}\psi_j^2$ Var $[\xi_t] = V[\varepsilon_t^a] < \infty$. Let $\gamma_j^c = \text{Cov}(\varepsilon_t^c, \varepsilon_{t-j}^c)$ and $\gamma_j^a = \text{Cov}(\varepsilon_t^a, \varepsilon_{t-j}^a)$. Then, $\gamma_j^c = \gamma_{cj}^a$. Thus, ε_t^c is stationary. Conversely, we assume that ε_t^a is not stationary. There is an integer, $d_a \geq 1$, such that ε_t^a ~ I (d_a). We can assume that $d_a = 1$ without loss of generality. Then, $\nabla\varepsilon_t^a$ is stationary, and it is trivial to prove that $\nabla\varepsilon_{ct}^a$ is also stationary. By Wold's decomposition, $\nabla\varepsilon_t^a = \psi(L)\kappa_t$, where $\psi(1) \neq 0$ and κ_t is white noise. Because $\nabla\varepsilon_t^c = \sum_{k=0}^{c-1}\nabla\varepsilon_{ct-k}^a = (1 + L + \cdots + L^{c-1})\ \psi(L)\kappa_{ct}$, $\nabla\varepsilon_t^c$ is stationary. Therefore, $\nabla\varepsilon_t^c = \theta(L)\eta_t$, where η_t is white noise with the same variance as κ_t and $\theta(L) = (1 + L + \cdots + L^{c-1})\psi(L)$. Thus, ε_t^c is an integrated process with order 1 because $\theta(1) \neq 0$. Consequently, ε_t^c is not stationary.

Theorem 1. Suppose ε_t is an integrated process. The following statements are equivalent.

(i) ε_t ~ I(d), (ii) ε_t^a ~ I(d), (iii) ε_t^c ~ I(d) for $d \geq 0$.

Proof. (i) is equivalent to (ii) by Lemma 1. (ii) holds if and only if (iii) holds for d = 0 by Lemma 2. Now, we consider the following equation:

$$\nabla^d \varepsilon_t^c = \sum_{k_1=0}^{c-1} \cdots \sum_{k_d=0}^{c-1} \nabla^d \varepsilon^a_{ct-k_1-\cdots-k_d}. \qquad (3)$$

Because ε_t is an integrated process, ε_t^a is an integrated process from the second equation in system (2), and equation (3) indicates that ε_t^c is also an integrated process. Then, by using equation (3) and Lemma 2, we can verify that $\nabla^d \varepsilon_t^a$ is stationary if any only if $\nabla^d \varepsilon_t^c$ is stationary for $d \geq 1$. Consequently, (ii) is equivalent to (iii) for a nonnegative integer *d*.

APPENDIX C: THE INFORMATION LOSS FUNCTION FOR ARC(3, 3)

Figure 1 shows the level curves of $\mathrm{ILF}(\phi_1, \phi_2, \phi_3)$ at some different values of ϕ_3. As ϕ_3 increases, the aggregation effects on the ILF decrease at a different rate according to (ϕ_1, ϕ_2). In particular, the ILF values decrease at a relatively much slower pace if $\theta < \pi / 2$ and $R \to 1$; thus, this allow the complex-type effects to gradually emerge in each of the level curves because ϕ_3 is close to 1.

Moreover, given ϕ_3, each layer of $\mathrm{ILF}(\phi_1, \phi_2, \phi_3)$ shares the same pattern as $\mathrm{ILF}(\phi_1, \phi_2)$, which indicates that the aggregation effects can explain the general shape of the level curves in Figure 1 as in $\mathrm{ARC}(2,3)$.

Figure 1 shows the direction of changes in the ILF of $\mathrm{ARC}(3,3)$. The level curves shift right to left as ϕ_3 increases, which implies that the brunt of forces to decrease the information loss occurs near $\phi_2 = 0$ along with $\phi_1 = k$. This result corresponds to the additive property of ILF.

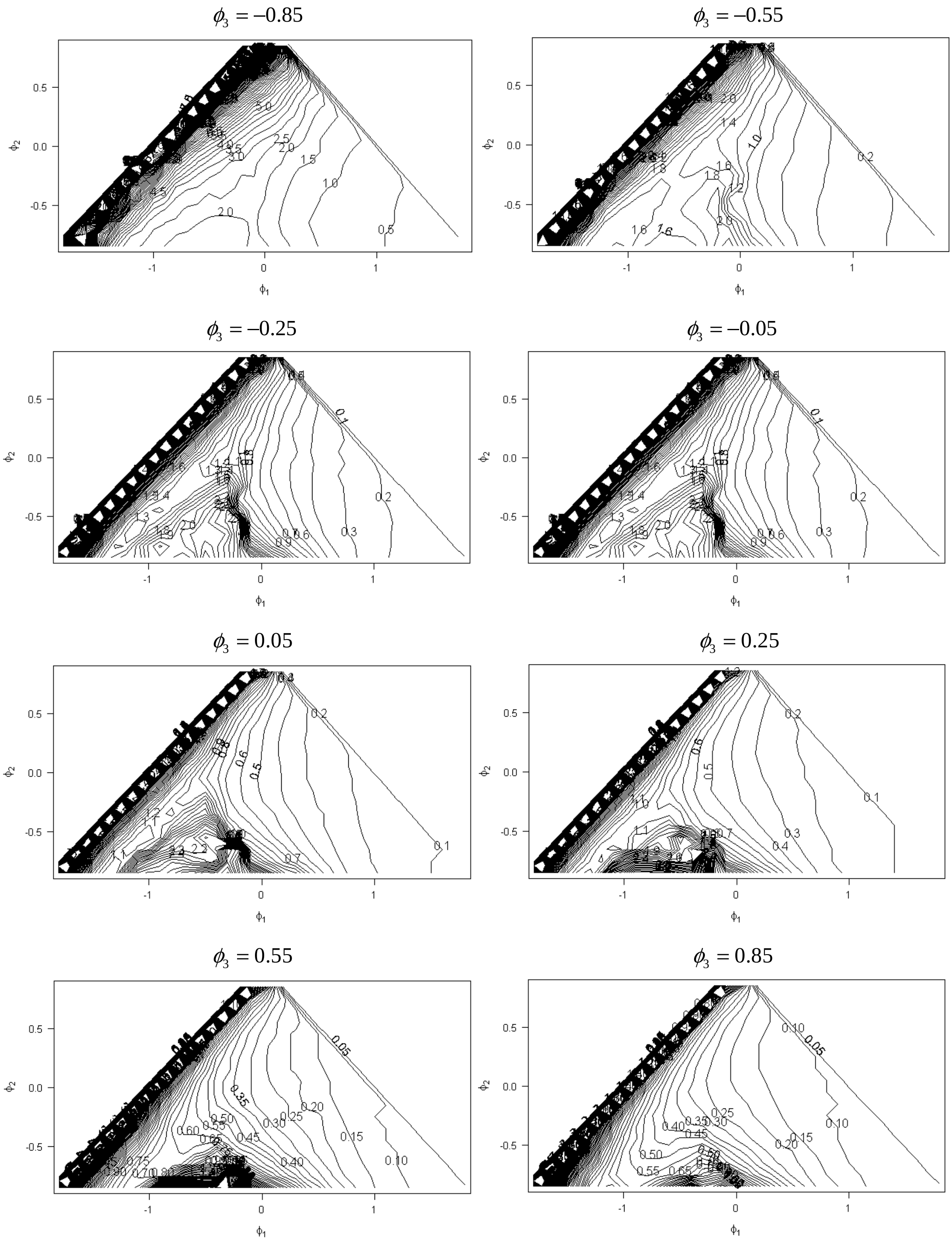

Figure 1. Monte Carlo simulation results for the ILF of ARC(3, 3). Each level curve has a similar pattern to the ILF of ARC(2, 3), which shows that the aggregation effects (Type I–III) can be applied to ARC(3, 3); the information loss function values are decreased along with the line, $\phi_1 = k$, as ϕ_3 increases. This confirms the additive property of the ILF.

APPENDIX D: INCONSISTENCY OF THE MLEs OF THE PARAMETERS IN ARC(1,2)

Consider a fraction of the time series, $\varepsilon_{i,t}$ and $\varepsilon_{i,t}^{c}$ for t = 1,…, 4, and i = 1,…, n, where

$\varepsilon_{i,1}^{c} = \varepsilon_{i,1} + \varepsilon_{i,2}$ and $\varepsilon_{i,2}^{c} = \varepsilon_{i,3} + \varepsilon_{i,4}$. Then,

$$\begin{pmatrix} \varepsilon_{i,1}^{c} \\ \varepsilon_{i,2}^{c} \end{pmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} \mu_i \\ \mu_i \end{bmatrix},\ \sigma_c^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right). \tag{4}$$

$$L(\mu_i, \sigma_c^2; \rho) = \left[\frac{1}{2\pi\sigma_c^2 \sqrt{(1-\rho^2)}} \right]^n \exp\left[-\frac{1}{2\sigma_c^2 (1-\rho^2)} \sum_{i=1}^{n} \left[\left(\varepsilon_{i,1}^{c} - \mu_i\right)^2 - 2\rho\left(\varepsilon_{i,1}^{c} - \mu_i\right)\left(\varepsilon_{i,2}^{c} - \mu_i\right) + \left(\varepsilon_{i,2}^{c} - \mu_i\right)^2 \right] \right],$$

where $\sigma_c^2 = Var(\varepsilon_{i,1}^{c}) = Var(\varepsilon_{i,2}^{c})$.

We assume that $-1 < \rho < 1$ is known. Then,

$$\frac{\partial \log L}{\partial \mu_i} = -\frac{1}{2\sigma_c^2(1-\rho^2)} \left[-2\left(\varepsilon_{i,1}^{c} - \mu_i\right) + 2\rho\left(\varepsilon_{i,1}^{c} + \varepsilon_{i,2}^{c} - 2\mu_i\right) - 2\left(\varepsilon_{i,2}^{c} - \mu_i\right) \right].$$

$$\frac{\partial \log L}{\partial \mu_i} = 0; \qquad \hat{\mu}_i = \frac{\varepsilon_{i,1}^{c} + \varepsilon_{i,2}^{c}}{2}.$$

Let $\alpha = \sigma_c^2 \sqrt{(1-\rho^2)}$. Then, $L(\hat{\mu}_i, \alpha; \rho) = \frac{1}{2\pi\alpha^n} \exp\left[-\frac{1}{2\alpha\sqrt{(1-\rho^2)}} \sum_{i=1}^{n} \frac{1+\rho}{2} \left(\varepsilon_{i,1}^{c} - \varepsilon_{i,2}^{c}\right)^2 \right].$

$$\frac{\partial \log(\hat{\mu}_i, \alpha; \rho)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{2\alpha^2\sqrt{(1-\rho^2)}} \sum_{i=1}^{n} \frac{1+\rho}{2} \left(\varepsilon_{i,1}^{c} - \varepsilon_{i,2}^{c}\right)^2.$$

$$\frac{\partial \log(\hat{\mu}_i, \alpha; \rho)}{\partial \alpha} = 0; \qquad \hat{\alpha} = \frac{1}{2n\sqrt{(1-\rho^2)}} \sum_{i=1}^{n} \frac{1+\rho}{2} \left(\varepsilon_{i,1}^{c} - \varepsilon_{i,2}^{c}\right)^2.$$ Then,

$$\hat{\sigma}_c^2 = \frac{1}{2n(1-\rho^2)} \sum_{i=1}^{n} \frac{1+\rho}{2} \left(\varepsilon_{i,1}^{c} - \varepsilon_{i,2}^{c}\right)^2. \tag{5}$$

From (4) $\varepsilon_{i,1}^{c} - \varepsilon_{i,2}^{c} \overset{iid}{\sim} N\left(0,\ 2(1-\rho)\sigma_c^2\right)$, and from (5), $\hat{\sigma}_c^2 \xrightarrow{p} \frac{1+\rho}{4(1-\rho^2)} \times 2(1-\rho)\sigma_c^2 = \frac{\sigma_c^2}{2} < \sigma_c^2$ by the weak law of large numbers. Thus, σ_c^2 converges in probability to a constant number that is less than σ_c^2.

In ARC(1, 2):

$$\varepsilon_t = \phi\varepsilon_{t-1} + \eta_t \qquad \text{for} \quad t = 1,\ldots,cT,$$
$$\varepsilon_t^c = \varepsilon_{2t} + \varepsilon_{2t-1} \qquad \text{for} \quad t = 1,\ldots,T,$$

where $\eta_t \sim NID(0,\sigma_\eta^2)$. $\sigma_c^2 = \mathrm{Var}(\varepsilon_t^c) = \mathrm{Var}(\varepsilon_{2t} + \varepsilon_{2t-1}) = 2(1+\phi)\mathrm{Var}(\varepsilon_t) = 2(1+\phi)\dfrac{\sigma_\eta^2}{1-\phi^2} = \dfrac{2\sigma_\eta^2}{1-\phi}$.

$\dfrac{\partial\sigma_c^2}{\partial\phi} = \dfrac{2\sigma_\eta^2}{(1-\phi)^2} > 0$ and $\dfrac{\partial\sigma_c^2}{\partial\sigma_\eta^2} = \dfrac{2}{1-\phi} > 0$, where $-1 < \phi < 1$; thus, σ_c^2 increases in ϕ and σ_η^2. Let $\sigma_c^2 = f(\phi)$, where $f \in C^\infty(-1,1)$ and $f' > 0$. Then, $\hat{\phi} = f^{-1}(\hat{\sigma}_c^2)$.

$$\hat{\phi} = f^{-1}(\hat{\sigma}_c^2) \xrightarrow{p} f^{-1}\left(\frac{\sigma_c^2}{2}\right) = \alpha_\phi f^{-1}\left(\sigma_c^2\right) < f^{-1}\left(\sigma_c^2\right) = \phi, \tag{6}$$

for some $0 < \alpha_\phi < 1$. Similarly, $\hat{\sigma}_\eta^2 \xrightarrow{p} \alpha_\eta\sigma_\eta^2$, for some $0 < \alpha_\eta < 1$. Consequently, in this fraction of the time series, the MLEs of ϕ and σ_η^2 converge to some constants that are less than the true values. This result aligns with Neyman and Scott's (1948) proposition that maximum likelihood estimates of the structural parameters related to a partially consistent series of observations need not be consistent. It is analytically intractable to consider the entire time series to show this phenomenon. Table 4 in the article, however, confirms that the Monte Carlo simulations provide the maximum likelihood estimates of ϕ and σ_η^2 that are less than the true values.

APPENDIX E: TEST RESULTS

Table 1. Unit root and cointegration tests: Real retail and food services sales (RFS), personal consumption expenditure (PCE), and unemployment rate (UER) of the United States (1992.1–2004.12)

	Unit Root Test			Cointegration Rank Test							
	Monthly	Quarterly		Monthly				Quarterly			
Series	Dickey–Fuller		No. of CEs[a]	Trace	Eigenvalue	Cointegrating Vectors		Trace	Eigenvalue	Cointegrating Vectors	
RFS	-2.579	-2.019	None	54.876^{*}	0.215^{*}	1.000	0.000	44.066^{*}	0.412^{*}	1.000	0.000
PCE	-0.579	-0.951	At most 1	17.908^{*}	0.108^{*}	0.000	1.000	18.590^{*}	0.265^{*}	0.000	1.000
UER	-1.081	-2.180	At most 2	0.454	0.003	-0.114 (0.225)	-0.039 (0.021)	3.827	0.077	-0.372 (0.204)	-0.062 (0.020)
	1	2	No. of lags	2				3			

NOTE: The asterisk signifies rejection of the corresponding null hypothesis at the 5% level of significance. [a] The number of cointegrating equations.

REFERENCES

Chow, G. C., and Lin, A. (1971), "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series," *Review of Economics and Statistics*, 53, 372-375.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59(6), 1551-1580.

Joo, Y., and Jun, D. (1997), "State Space Trend-Cycle Decomposition of the ARIMA(1,1,1) Process," *Journal of Forecasting*, 16, 411-424.

Jun, D. B., Kim, D. S., Park, S., and Park, M. H. (2012), "Parameter Space Restrictions in State Space Models," *Journal of Forecasting*, 31, 109-123.

Luenberger, D. G. (1979), *Introduction to Dynamic Systems: Theory, Models, and Applications*, New York: John Wiley & Sons.

Neyman, J., and Scott, E. L. (1948). "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 16, 1-32.