TESTING FOR A UNIT ROOT IN LEE–CARTER MORTALITY MODEL

BY

XUAN LENG AND LIANG PENG

ABSTRACT

Motivated by a recent discovery that the two-step inference for the Lee–Carter mortality model may be inconsistent when the mortality index does not follow from a nearly integrated AR(1) process, we propose a test for a unit root in a Lee–Carter model with an AR(p) process for the mortality index. Although testing for a unit root has been studied extensively in econometrics, the method and asymptotic results developed in this paper are unconventional. Unlike a blind application of existing R packages for implementing the two-step inference procedure in Lee and Carter (1992) to the U.S. mortality rate data, the proposed test rejects the null hypothesis that the mortality index follows from a unit root AR(1) process, which calls for serious attention on using the future mortality projections based on the Lee–Carter model in policy making, pricing annuities and hedging longevity risk. A simulation study is conducted to examine the finite sample behavior of the proposed test too.

KEYWORDS

AR process, Lee–Carter model, unit root.

1. INTRODUCTION

The increased life expectancy has posted serious challenges to insurance companies and pension funds for managing longevity risk. For hedging longevity risk and pricing annuities, mortality models and their inferences play an important role (see Frees *et al.,* 1996; Currie *et al.,* 2004; Cairns *et al.,* 2008; Njenga and Sherris 2011). Although many types of mortality models have been proposed in the literature of actuarial science (see Haberman and Renshaw, 2008; Chen and Cox, 2009; Bauer *et al.,* 2012; Yang and Wang, 2013), a quite popular model is the so-called Le–Carter model, where Lee and Carter (1992) proposed to model the logarithm of the central mortality rate by

$$\log m_{x,t} = \alpha_x + \beta_x k_t + \epsilon_{x,t},$$

Astin Bulletin 47(3), 715-735. doi: 10.1017/asb.2017.24

where $m_{x,t}$ denotes the central mortality rate for age group $x = 1, \ldots, K$ at time period $t = 1, \ldots, T$ and $\epsilon_{x,t}$'s are random errors. Here, k_t is an unobserved factor called mortality index, α_x and β_x are unknown parameters. Due to identification issue, conditions $\sum_{x=1}^{K} \beta_x = 1$ and $\sum_{t=1}^{T} k_t = 0$ are imposed in finding estimators $\hat{\alpha}_x$, $\hat{\beta}_x$, $\hat{k}_t$, which are obtained by the singular value decomposition method. A recent development on singular value decomposition method for analyzing mortality data is given in Zhang *et al.* (2013). For effective prediction, Lee and Carter (1992) further proposed to fit an ARIMA(p, d, q) model to $\hat{k}_t$'s, and researchers found that an ARIMA$(p, 1, q)$ model is preferred in fitting existing mortality data.

Since the seminal paper of Lee and Carter (1992), many extensions and applications have appeared in the literature with an open R package (demography); see Brouhns *et al.* (2002), Li and Lee (2005), Girosi and King (2007), Cairns *et al.* (2011), Bisetti and Favero (2014), D'Amato *et al.* (2014) and Lin *et al.* (2014).

Although there are wide applications of the Lee–Carter model and its extensions in policy making, there is almost no discussion on whether the two-step inference procedure could lead to a correct identification of the dynamic structure of k_t, which plays a key role in forecasting mortality rates and managing longevity risks. Recently, Leng and Peng (2016) unfortunately proved that the two-step inference procedure in Lee and Carter (1992) may lead to a wrong identification of the dynamics of k_t if it is not a nearly integrated AR(1) model. Therefore, this raises an interesting question on how to test unit root in the Lee–Carter mortality model, where the dynamics of mortality index follows from an AR(p) model.

Since there exist no asymptotic results for the two-step inference procedure in Lee and Carter (1992), let us blindly apply the method to the U.S. mortality rate data from year 1933 to year 2010. More specifically, we consider Group I with age groups 0, 1–4, 5–9, ..., 105–109, 110+, and Group II with age groups 10–14, 15–19, ..., 65–69. We first employ the "demography" package in R to obtain estimators $\hat{k}_t$ with and without re-estimation, and then we use "lm" in R to fit both an AR(1) model and an AR(2) model with an intercept to the obtained $\hat{k}_t$'s. Re-estimation involves a step to match the number of deaths; see the "demography" package document for details. The results are reported in Tables 1 and 2.

By looking at the reported standard deviations, we may conclude that (i) a unit root AR(1) model is preferred for Group I; (ii) a unit root AR(1) model is preferred for Group II if the re-estimation is employed. That is, one may think the two-step inference procedure in Lee and Carter (1992) is applicable to these data sets since a unit root AR(1) model for the mortality index is preferred. Apparently, the reported standard deviations are inaccurate since they ignore the randomness in obtaining $\hat{k}_t$'s. Therefore, it is necessary to derive a formal test for testing whether the mortality index follows from a unit root AR(1) process. Although testing for a unit root has been studied extensively in the literature

TABLE 1

ESTIMATORS FOR COEFFICIENTS AND THEIR STANDARD DEVIATION FOR GROUP I. WE REPORT ESTIMATORS FOR COEFFICIENTS AND THEIR STANDARD DEVIATION IN BRACKETS BY USING THE "lm" IN SOFTWARE R TO FIT EITHER AN AR(1) MODEL OR AN AR(2) MODEL WITH AN INTERCEPT TO THE ESTIMATED k_t'S WITH AND WITHOUT RE-ESTIMATION OBTAINED BY THE R PACKAGE "DEMOGRAPHY" FOR U.S. MORTALITY DATA FROM YEAR 1933 TO YEAR 2010.

	Group I			
Model	Estimated k_t's	ϕ_0	ϕ_1	ϕ_2
AR(1)	No Re-estimation	−0.3265(0.0448)	0.9918(0.0064)	NA
	Re-estimation	−0.3366(0.0535)	1.0003(0.0075)	NA
AR(2)	No Re-estimation	−0.2983(0.0572)	1.1083(0.1131)	−0.1188(0.1123)
	Re-estimation	−0.3987(0.0652)	0.8443(0.1132)	0.1540(0.1135)

TABLE 2

ESTIMATORS FOR COEFFICIENTS AND THEIR STANDARD DEVIATION FOR GROUP II. WE REPORT ESTIMATORS FOR COEFFICIENTS AND THEIR STANDARD DEVIATION IN BRACKETS BY USING THE "lm" IN SOFTWARE R TO FIT EITHER AN AR(1) MODEL OR AN AR(2) MODEL WITH AN INTERCEPT TO THE ESTIMATED k_t'S WITH AND WITHOUT RE-ESTIMATION OBTAINED BY THE R PACKAGE "DEMOGRAPHY" FOR U.S. MORTALITY DATA FROM YEAR 1933 TO YEAR 2010.

	Group II			
Model	Estimated k_t's	ϕ_0	ϕ_1	ϕ_2
AR(1)	No Re-estimation	−0.2016(0.0324)	0.9847(0.0078)	NA
	Re-estimation	−0.2002(0.0295)	0.9967(0.0070)	NA
AR(2)	No Re-estimation	−0.1410(0.0372)	1.3178(0.1084)	−0.3325(0.1068)
	Re-estimation	−0.1915(0.0374)	1.0637(0.1150)	−0.0693(0.1148)

of econometrics, see Phillips and Perron (1988) and the recent review paper by Xiao (2014), methods and asymptotic results developed in this paper are quite different from existing ones due to the special structure of the mortality model; see Section 2 for details.

We organize this paper as follows. Section 2 presents the model, method and asymptotic results for testing a unit root. A simulation study and real data analysis are given in Section 3, which indicates that the proposed test requires a large number of observations ($\geq$ 100) in order to have a reasonable size. Since practitioners usually use a short period of mortality data, it is important to develop a more powerful test in future for ensuring the applicability of the widely employed Lee–Carter model and its extensions. Some conclusions are summarized in Section 4. All proofs are put in the Appendix.

2. MODEL, METHODOLOGY AND ASYMPTOTIC RESULTS

Consider the following Lee–Carter model:

$$\log m_{x,t} = \alpha_x + \beta_x k_t + \epsilon_{x,t}, \quad k_t = \phi_0 + \phi_1 k_{t-1} + \phi_2 k_{t-2} + e_t \tag{1}$$

for $x = 1, \ldots, K$ and $t = 1, \ldots, T$, where $\epsilon_{x,t}$'s and e_t's are independent random errors with $\mathsf{E}\,\epsilon_{x,t} = \mathsf{E}\,e_t = 0$, $\mathsf{E}\,\epsilon_{x,t}^2 = \sigma_x^2$ and $\mathsf{E}\,e_t^2 = \sigma^2$. As shown in Leng and Peng (2016), the two-step estimation procedure in Lee and Carter (1992) may be inconsistent when $\{k_t\}$ does not follow from a nearly integrated AR(1) model. Therefore, an interesting question is how to test $H_0 : \phi_1 = 1 \,\&\, \phi_2 = 0$ for the above Lee–Carter model. Note that k_t's are unobserved and the two-step inference procedure in Lee and Carter (1992) cannot be employed due to its inconsistency and unknown asymptotic behavior.

Rewrite (1) as

$$\log m_{x,t} = \delta_x + \phi_1 \log m_{x,t-1} + \phi_2 \log m_{x,t-2} + u_{x,t}, \tag{2}$$

where

$$u_{x,t} = \beta_x e_t + \epsilon_{x,t} - \phi_1 \epsilon_{x,t-1} - \phi_2 \epsilon_{x,t-2}, \tag{3}$$

and $\delta_x = (1 - \sum_{s=1}^2 \phi_s)\alpha_x + \phi_0 \beta_x$, which is independent of t and represents the age-specific trend. Put

$$y_{x,t}^{(i)} = \log m_{x,t-i} - T^{-1} \sum_{j=1}^{T} \log m_{x,j-i} \quad \text{and} \quad u_{x,t}^* = u_{x,t} - \frac{1}{T} \sum_{s=1}^{T} u_{x,s},$$

where $\log m_{x,t}$ is defined to be zero for $t \leq 0$ and $x = 1, \ldots, K$. Therefore, (2) is equivalent to

$$y_{x,t}^{(0)} = \phi_1 y_{x,t}^{(1)} + \phi_2 y_{x,t}^{(2)} + u_{x,t}^*. \tag{4}$$

This motivates us to estimate ϕ_1 and ϕ_2 by minimizing the following least squares:

$$\sum_{x=1}^{K} \sum_{t=1}^{T} \left(y_{x,t}^{(0)} - \phi_1 y_{x,t}^{(1)} - \phi_2 y_{x,t}^{(2)} \right)^2,$$

which leads to

$$\tilde{\phi}_1 = \frac{D_{0,1} D_{2,2} - D_{0,2} D_{1,2}}{D_{1,1} D_{2,2} - D_{1,2}^2} \quad \text{and} \quad \tilde{\phi}_2 = \frac{D_{1,1} D_{0,2} - D_{0,1} D_{1,2}}{D_{1,1} D_{2,2} - D_{1,2}^2},$$

with $D_{i,j} = \sum_{x=1}^{K} \sum_{t=1}^{T} y_{x,t}^{(i)} y_{x,t}^{(j)}$.

Although testing for a unit root has received extensive studies in the literature of econometrics, the study here is quite different from existing ones. First, under $H_0 : \phi_1 = 1 \,\&\, \phi_2 = 0$, $\log m_{x,t-1}$ is correlated with $u_{x,t}$ and thus the above least squares estimators are biased. Second, it follows from (A21) in the proof of Theorem 1 that the difference of the following two terms from the score equations,

$$T^{-3/2}\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)}\left(y_{x,t}^{(0)}-\phi_1 y_{x,t}^{(1)}-\phi_2 y_{x,t}^{(2)}\right) \quad \text{and}$$

$$T^{-3/2}\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(2)}\left(y_{x,t}^{(0)}-\phi_1 y_{x,t}^{(1)}-\phi_2 y_{x,t}^{(2)}\right),$$

converges in probability to zero as $T \to \infty$ when $H_0 : \phi_1 = 1 \& \phi_2 = 0$ and there exists a non-zero trend (i.e., $\sum_{x=1}^{K} \delta_x^2 > 0$), which means that the joint asymptotic distribution of the above two terms is degenerate. Similarly, it follows from (A10) and (A12) in the proof of Theorem 1 that the difference of the following two terms from the score equations,

$$T^{-1}\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)}\left(y_{x,t}^{(0)}-\phi_1 y_{x,t}^{(1)}-\phi_2 y_{x,t}^{(2)}\right) \quad \text{and}$$

$$T^{-1}\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(2)}\left(y_{x,t}^{(0)}-\phi_1 y_{x,t}^{(1)}-\phi_2 y_{x,t}^{(2)}\right),$$

converges in probability to a constant as $T \to \infty$ when $H_0 : \phi_1 = 1 \& \phi_2 = 0$ and there exists no trend (i.e., $\sum_{x=1}^{K} \delta_x^2 = 0$), which means that the joint asymptotic distribution is degenerate. To overcome the second issue of having a degenerate limiting distribution, we propose to test $H_0 : \phi_1 = 1 \& \phi_1 + \phi_2 = 1$, i.e., to use estimators for ϕ_1 and $\phi_1 + \phi_2$, which has some similarity to the idea of the augmented Dickey–Fuller test (see Said and Dickey, 1984). For dealing with the first issue, a well-known technique developed and commonly employed in econometrics is the so-called instrumental variable method. Due to the special structure of $u_{x,t}$, finding an instrumental variable is not easy at all. Instead we propose the following bias corrected least squares estimator for ϕ_1.

Write

$$\begin{aligned} &D_{0,1}D_{2,2} - D_{0,2}D_{1,2} - (D_{1,1}D_{2,2} - D_{1,2}^2) \\ &= D_{2,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\right] - D_{1,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(2)}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\right] \\ &= -\left[\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\right]\left[\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(2)}\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right] \end{aligned}$$

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

$$+D_{1,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right]$$

$$= I_1 + I_2,$$

$$\begin{aligned}
&D_{1,1}D_{2,2}-D_{1,2}^2\\
&\quad=(D_{1,1}-D_{1,2})(D_{2,2}-D_{1,2})+D_{1,2}[D_{1,1}-D_{1,2}-(D_{1,2}-D_{2,2})]\\
&\quad=-\left[\sum_{x=1}^{K}\sum_{t=1}^{T}y_{x,t}^{(1)}\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right]\left[\sum_{x=1}^{K}\sum_{t=1}^{T}y_{x,t}^{(2)}\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right]\\
&\qquad+D_{1,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T}\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)^2\right]\\
&\quad=II_1+II_2,
\end{aligned}$$

and

$$\begin{aligned}
&D_{1,1}D_{0,2}-D_{0,1}D_{1,2}\\
&\quad=D_{1,1}(D_{0,2}-D_{1,2})-D_{1,2}(D_{0,1}-D_{1,1})\\
&\quad=\left[\sum_{x=1}^{K}\sum_{t=1}^{T}y_{x,t}^{(1)}\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right]\left[\sum_{x=1}^{K}\sum_{t=1}^{T}y_{x,t}^{(2)}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\right]\\
&\qquad-D_{1,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right]\\
&\quad=III_1-I_2.
\end{aligned}$$

Since $\mathsf{E}\,(u_{x,t}u_{x,t-1})\neq 0$, it follows from (A10)–(A12) and (A21) in the proof of Theorem 1 that I_2 and II_2 have the same order and dominate I_1, II_1 and III_1 under $H_0:\phi_1=1\,\&\,\phi_2=0$. When we say I_2 dominates I_1, it means $I_1=o_p(I_2)$ as $T\to\infty$. Therefore, the inconsistency of the least squares estimators is due to term I_2, and an obvious bias corrected estimator for ϕ_1 is

$$\tilde{\phi}_1-\frac{D_{1,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T}\left(y_{x,t}^{(0)}-y_{x,t}^{(1)}\right)\left(y_{x,t}^{(1)}-y_{x,t}^{(2)}\right)\right]}{D_{1,1}D_{2,2}-D_{1,2}^2}.$$

However, after writing the term $\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)}(y_{x,t}^{(0)} - y_{x,t}^{(1)})$ in I_1 as

$$\sum_{x=1}^{K}\sum_{t=1}^{T} u_{x,t}^{*} y_{x,t}^{(1)} + (\phi_1 - 1)\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)}\left(y_{x,t}^{(1)} - y_{x,t}^{(2)}\right) + (\phi_1 + \phi_2 - 1)\sum_{x=1}^{K}\sum_{t=1}^{T} y_{x,t}^{(1)} y_{x,t}^{(2)},$$

which has the same limit as $\sum_{x=1}^{K}\sum_{t=1}^{T} u_{x,t}^{*} y_{x,t}^{(1)}$ when $\phi_1 + \phi_2 - 1 = 0$ and $\phi_1 - 1 = o(1)$, we conclude that the above bias corrected estimator cannot detect the case of $\phi_1 - 1 = o(1)$ when $\phi_2 + \phi_1 - 1 = 0$. To better detect the departure of ϕ_1 from one, we propose the following bias corrected estimator for ϕ_1:

$$\hat{\phi}_1 = \tilde{\phi}_1 - \frac{D_{1,2}\left[\sum_{x=1}^{K}\sum_{t=1}^{T}(y_{x,t}^{(0)} - y_{x,t}^{(1)})(y_{x,t}^{(1)} - y_{x,t-1}^{(2)})\right]}{D_{1,1}D_{2,2} - D_{1,2}^{2}}.$$

The difference from the previous bias-corrected estimator is that we use $y_{x,t-1}^{(2)}$ instead of $y_{x,t}^{(2)}$, which turns out to be quite effective in detecting the departure of ϕ_1 from one both theoretically and practically. Note that under $H_0 : \phi_1 = 1 \,\&\, \phi_2 = 0$, $\tilde{\phi}_1 + \tilde{\phi}_2$ converges in probability to one since the I_2 term disappears. In conclusion, we propose to consider the joint limit of estimators $\hat{\phi}_1 - 1$ and $\tilde{\phi}_1 + \tilde{\phi}_2 - 1$ for testing $H_0 : \phi_1 = 1 \,\&\, \phi_1 + \phi_2 = 1$, which is equivalent to $H_0 : \phi_1 = 1 \,\&\, \phi_2 = 0$. Throughout we assume that

C1) $\{\epsilon_t = (\epsilon_{1,t}, \ldots, \epsilon_{K,t})^{\tau}\}_{t=1}^{T}$ is a sequence of independent and identically distributed random vectors with zero means and covariance matrix $\Sigma^{\epsilon} = (\sigma_{i,j}^{\epsilon})$, $\{e_t\}_{t=1}^{T}$ is a sequence of independent and identically distributed random variables with zero mean and finite variance σ^2, and these two sequences are independent. Here, A^{τ} denotes the transpose of the vector or matrix A;

C2) $\mathsf{E}(||\epsilon_t||^{\eta}) + \mathsf{E}(|e_t|^{\eta}) < \infty$ for some $\eta > 2$.

Put $u_t = (u_{1,t}, \ldots, u_{K,t})^{\tau}$ and it is easy to show that, under conditions of Theorem 1 below,

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\mathsf{E}\left\{\sum_{t=1}^{T} u_{i,t}\sum_{t=1}^{T} u_{j,t}\right\} \\
&= \lim_{T\to\infty}\frac{1}{T}\left\{\sum_{t=1}^{T}\mathsf{E}(u_{i,t}u_{j,t}) + 2\sum_{t=1}^{T}\mathsf{E}(u_{i,t}u_{j,t-1})\right\} \\
&= \beta_i\beta_j\sigma^2 \\
&=: \sigma_{i,j}.
\end{aligned}$$

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

Define $\Sigma = (\sigma_{i,j})_{1\le i,j\le K}$ and

$$X_T(r) = \frac{1}{\sqrt{T}}\Sigma^{-1/2}\left(\sum_{t=1}^{[r_1T]} u_{1,t},\ldots,\sum_{t=1}^{[r_KT]} u_{K,t}\right)^\tau \quad \text{for} \quad r = (r_1,\ldots,r_K)^\tau \in [0,1]^K.$$

Then, like the proofs in Phillips and Durlauf (1986), we have

$$X_T(r) \xrightarrow{D} W(r) = \big(W_1(r_1),\ldots,W_K(r_K)\big)^\tau, \tag{5}$$

where "$\xrightarrow{D}$" denotes convergence in space $D([0,1]^K)$ and $W_1(r_1),\ldots,W_K(r_K)$ are independent Brownian motions. When $r_1 = \cdots = r_K = s$ for some $s \in [0,1]$, we simply write $X_T(r)$ as $X_T(s)$.

Throughout define $\delta = (\delta_1,\ldots,\delta_K)^\tau$, $\log m_t = (\log m_{1,t},\ldots,\log m_{K,t})^\tau$, $u_t = (u_{1,t},\ldots,u_{K,t})^\tau$, $y_t^{(i)} = \left(y_{1,t}^{(i)},\ldots,y_{K,t}^{(i)}\right)^\tau$, $J_d(s) = W(s) + d\int_0^s e^{d(s-t)}W(t)\,dt$ and $\tilde{J}_d(s) = J_d(s) - \int_0^1 J_d(t)\,dt$. We use "$\xrightarrow{d}$" and "$\xrightarrow{p}$" to denote convergence in distribution and convergence in probability, respectively.

Theorem 1. *Suppose model (1) holds with conditions C1) and C2).*

(i) *If $\delta^\tau\delta = 0$, i.e., $\delta \equiv 0$, then for $\phi_1 = 1 + d_1/\sqrt{T}$ & $\phi_1 + \phi_2 = 1 + d_2/T$, we have*

$$\big(\sqrt{T}(\hat\phi_1 - 1),\, T(\tilde\phi_1 + \tilde\phi_2 - 1)\big) \xrightarrow{d} (Z_1, \tilde Z_1),$$

where

$$Z_1 \sim N\left(\frac{d_1\sum_{x=1}^K \sigma_{x,x}^\epsilon}{\sum_{x=1}^K(\sigma_{x,x}+2\sigma_{x,x}^\epsilon)}, \frac{\sum_{x_1=1}^K\sum_{x_2=1}^K\big((\sigma_{x_1,x_2}+2\sigma_{x_1,x_2}^\epsilon)^2 + 2(\sigma_{x_1,x_2}^\epsilon)^2\big)}{\big(\sum_{x=1}^K(\sigma_{x,x}+2\sigma_{x,x}^\epsilon)\big)^2}\right),$$

$$\tilde Z_1 = \frac{Z_2\sum_{x=1}^K\sigma_{x,x}^\epsilon}{Z_3\sum_{x=1}^K(\sigma_{x,x}+2\sigma_{x,x}^\epsilon)} + \frac{Z_2 + \sum_{x=1}^K\sigma_{x,x}^\epsilon}{Z_3} + d_2\frac{\sum_{x=1}^K(\sigma_{x,x}+3\sigma_{x,x}^\epsilon)}{\sum_{x=1}^K(\sigma_{x,x}+2\sigma_{x,x}^\epsilon)},$$

$$Z_2 = \frac{1}{2}\mathrm{tr}\left(\Sigma\big(\tilde J_{d_2}(1)\tilde J_{d_2}^\tau(1) - \tilde J_{d_2}(0)\tilde J_{d_2}^\tau(0) - 2(d_2 + d_1^2)\right.$$
$$\left.\int_0^1 \tilde J_{d_2}(s)\tilde J_{d_2}^\tau(s)\,ds\big) - \Sigma - 2\Sigma^\epsilon\right),$$

$Z_3 = \mathrm{tr}\left(\Sigma\int_0^1\tilde J_{d_2}(s)\tilde J_{d_2}^\tau(s)\,ds\right)$ *and Z_1 is independent of $W(s)$. Here,* $\mathrm{tr}(A)$ *denotes the trace of matrix A.*

(ii) *If $\delta^\tau\delta > 0$, i.e., $\delta \not\equiv 0$, then for $\phi_1 = 1 + d_1/\sqrt{T}$ & $\phi_1 + \phi_2 = 1 + d_2/T^{3/2}$, we have*

$$\big(\sqrt{T}(\hat\phi_1 - 1),\, T^{3/2}(\tilde\phi_1 + \tilde\phi_2 - 1)\big) \xrightarrow{d} (Z_1, \tilde Z_1^*),$$

where

$$\tilde{Z}_1^* = \frac{12\sum_{x=1}^K(\sigma_{x,x}+3\sigma_{x,x}^\epsilon)}{\delta^\tau\delta\sum_{x=1}^K(\sigma_{x,x}+2\sigma_{x,x}^\epsilon)}Z_4 + d_2\frac{\sum_{x=1}^K(\sigma_{x,x}+3\sigma_{x,x}^\epsilon)}{\sum_{x=1}^K(\sigma_{x,x}+2\sigma_{x,x}^\epsilon)}$$

and

$$Z_4 = \mathrm{tr}\left(\Sigma^{1/2}\left(\frac{W(1)}{2} - \int_0^1 W(s)\,ds\right)\delta^\tau\right) \sim N\left(0, \frac{1}{12}\mathrm{tr}(\Sigma\delta\delta^\tau)\right).$$

Based on the above theorem, we can develop a test for a unit root in the Lee–Carter model via estimating the unknown quantities in the limit. For mortality rates, one usually has $\delta^T\delta > 0$ in practice. Hence, we apply the asymptotic results in Theorem 1 (ii) to test a unit root for the Lee–Carter model as follows.

Under $H_0 : \phi_1 = 1 \,\&\, \phi_2 = 0$, we have from (2) that

$$\log m_{x,t} = \delta_x + \log m_{x,t-1} + u_{x,t},\ \ x = 1, \ldots, K.$$

Since $\mathsf{E}\,(u_{x,t}) = 0$, we estimate δ_x by

$$\hat{\delta}_x = T^{-1}\sum_{t=1}^T(\log m_{x,t} - \log m_{x,t-1}) = T^{-1}\log m_{x,T}.$$

Define

$$\hat{\delta} = (\hat{\delta}_1, \ldots, \hat{\delta}_K)^\tau, \quad \hat{u}_t^* = (\hat{u}_{1,t}^*, \ldots, \hat{u}_{K,t}^*)^\tau,$$

$$\hat{u}_{x,t}^* = y_{x,t}^{(0)} - \hat{\phi}_1\left(y_{x,t}^{(1)} - y_{x,t}^{(2)}\right) - (\tilde{\phi}_1 + \tilde{\phi}_2)y_{x,t}^{(2)},$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^T\hat{u}_t^*\hat{u}_t^{*\tau} + \frac{2}{T}\sum_{t=1}^T\hat{u}_t^*\hat{u}_{t-1}^{*\tau}, \quad \hat{\Delta}_1 = \mathrm{tr}(\hat{\Sigma}), \quad \hat{\Sigma}^\epsilon = -\frac{1}{T}\sum_{t=1}^T\hat{u}_t^*\hat{u}_{t-1}^{*\tau},$$

$$\hat{\Delta}_2 = \mathrm{tr}(\hat{\Sigma}^\epsilon) \quad \text{and} \quad \hat{\Delta}_3 = \mathrm{sum}\big((\hat{\Sigma} + 2\hat{\Sigma}^\epsilon)^2 + 2(\hat{\Sigma}^\epsilon)^2\big),$$

where A^2 denotes the matrix with each element being the square of the corresponding element in the matrix A and sum(A) denotes the summation of all elements in the matrix A. Then under $H_0 : \phi_1 = 1 \,\&\, \phi_2 = 0$, our test statistic is defined as

$$Z = \frac{T(\hat{\phi}_1 - 1)^2(\hat{\Delta}_1 + 2\hat{\Delta}_2)^2}{\hat{\Delta}_3} + \frac{T^3(\tilde{\phi}_1 + \tilde{\phi}_2 - 1)^2(\hat{\delta}^\tau\hat{\delta})^2(\hat{\Delta}_1 + 2\hat{\Delta}_2)^2}{12\,\mathrm{tr}(\hat{\Sigma}\hat{\delta}\hat{\delta}^\tau)\,(\hat{\Delta}_1 + 3\hat{\Delta}_2)^2}. \tag{6}$$

It immediately follows from Theorem 1 (ii) that Z has a chi-squared limiting distribution with two degrees of freedom as $T \to \infty$ when $\delta \not\equiv 0$. Hence, in this case, we reject H_0 at level α if $Z > \chi^2_{1-\alpha,2}$, where $\chi^2_{1-\alpha,2}$ denotes the $(1-\alpha)$-th quantile of a chi-squared distribution with two degrees of freedom.

TABLE 3

U.S. MORTALITY RATES. WE REPORT ESTIMATORS $\hat{\phi}_1$, $\tilde{\phi}_1+\tilde{\phi}_2$, $\hat{\delta}^\tau\hat{\delta}$ FOR ϕ_1, $\phi_1+\phi_2$ AND $\delta^\tau\delta$, RESPECTIVELY. THE COMPUTED TEST STATISTIC Z DEFINED IN (6) AND ITS P-VALUE ARE REPORTED TOO. GROUP I MEANS AGE GROUPS 0, 1–4, 5–9, …, 105–109, 110+, AND GROUP II MEANS AGE GROUPS 10–14, …, 65–69.

	Group I	Group II
$\hat{\phi}_1$	1.0301	1.0674
$\tilde{\phi}_1+\tilde{\phi}_2$	0.9519	0.9829
$\hat{\delta}^\tau\hat{\delta}$	0.1161	0.0800
Z	533	148
P-value	0.0000	0.0000

Remark 1. *If the mortality index follows from an AR(p) model:*

$$k_t=\phi_0+\sum_{s=1}^{p}\phi_s k_{t-s}+e_t,$$

then, using the same notation as above, we can write

$$y_{x,t}^{(0)}=\sum_{s=1}^{p}\phi_s y_{x,t}^{(s)}+u_{x,t}^*=\phi_1(y_{x,t}^{(1)}-y_{x,t}^{(2)}-\cdots-y_{x,t}^{(p)})+\sum_{s=2}^{p}(\phi_1+\phi_s)y_{x,t}^{(s)}+u_{x,t}^*,$$

where $u_{x,t}^*=u_{x,t}-\frac{1}{T}\sum_{s=1}^{T}u_{x,s}$ *with* $u_{x,t}=\beta_x e_t+\epsilon_{x,t}-\sum_{s=1}^{p}\phi_s\epsilon_{x,t-s}$. *Based on the least squares estimators for* $\phi_1, \phi_1+\phi_2, \dots, \phi_1+\phi_p$, *a similar test to the case of* $p=2$ *can be derived for testing* $H_0:\phi_1=1=\phi_1+\phi_2=\cdots=\phi_1+\phi_p$, *which is equivalent to* $H_0:\phi_1=1, \phi_2=\cdots=\phi_p=0$. *Here, we skip the details since mortality rate data usually do not prefer* $p>2$.

3. DATA ANALYSIS AND SIMULATION STUDY

First, we apply the proposed test in Section 2 for testing $H_0:\phi_1=1\,\&\,\phi_2=0$ in model (1) to the U.S. mortality rate data from year 1933 to year 2010 with two different age groups by assuming $\delta^T\delta>0$. Group I contains $K=24$ age groups, which are 0, 1–4, 5–9, 10–14, …, 105–109, 110+, and Group II represents $K=12$ age groups, which are 10–14, 15–19, …, 65–69. This data set is available in the Human Mortality Database, http://www.mortality.org. Table 3 reports $\hat{\phi}_1$, $\tilde{\phi}_1+\tilde{\phi}_2$, $\hat{\delta}^\tau\hat{\delta}$, the test statistic Z given in (6) and its P-value. We also plot $\hat{\delta}$, diag$(\hat{\Sigma})$ and diag$(\hat{\Sigma}^\epsilon)$ in Figure 1, which are employed to set up the following simulation study. Here, diag(A) means replacing the off-diagonal elements of matrix A by 0.

In contradictory to the blind application of the two-step inference procedure in Lee and Carter (1992) given Section 1, the P-values in Table 3 clearly reject the null hypothesis $H_0:\phi_1=1$ and $\phi_2=0$. Hence, applying the two-step inference procedure in Lee and Carter (1992) to the U.S. mortality rate data is

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

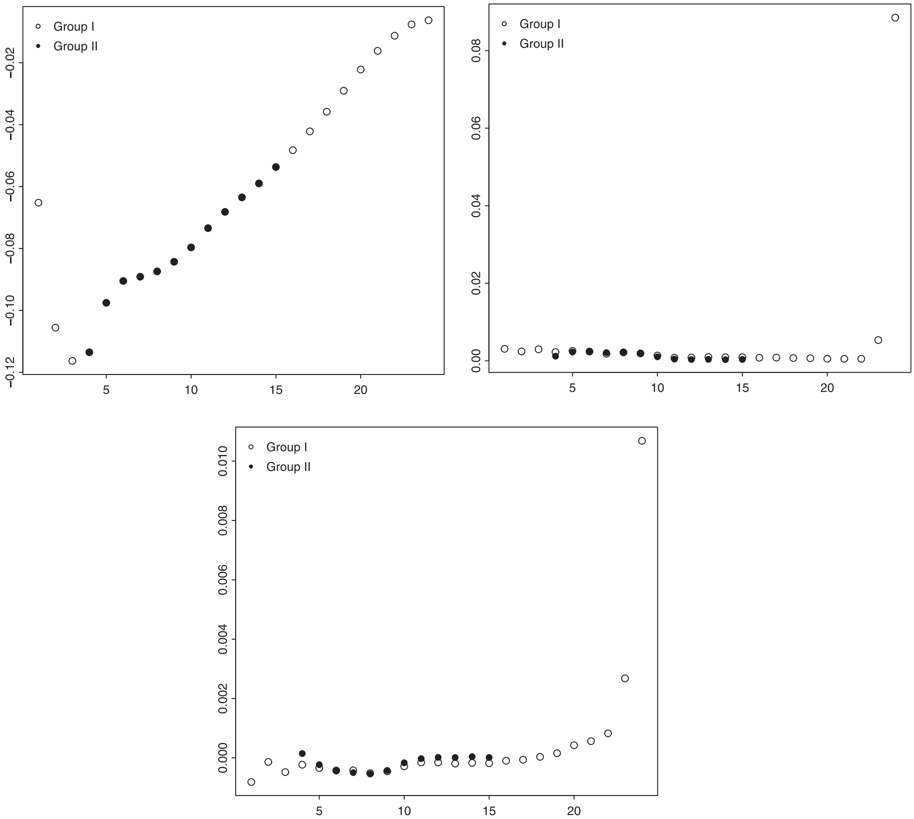

FIGURE 1: Estimates $\hat{\delta}$, diag($\hat{\Sigma}$) and diag($\hat{\Sigma}^{\epsilon}$) are plotted for the U.S. mortality rates.

problematic, and one should be cautious to use the future mortality projections based on the Lee–Carter model in policy making, pricing annuities and hedging longevity risk.

Next, we examine the finite sample performance of the proposed test by generating observations from (2) and (3) with $\phi_1 = 1 + d/\sqrt{T}$, $\phi_1 + \phi_2 = 1 + d/T^{3/2}$, $\delta = \hat{\delta}$, $(\beta_1, \ldots, \beta_K)^{\tau} e_t \sim (\sqrt{\hat{\sigma}_{1,1}}, \ldots, \sqrt{\hat{\sigma}_{K,K}})^{\tau} N(0, 1)$ and $\epsilon_t \sim N\big(0, |\mathrm{diag}(\hat{\Sigma}^{\epsilon})|\big)$, where $\hat{\delta}$, $\hat{\Sigma} = (\hat{\sigma}_{i,j})$ and $\hat{\Sigma}^{\epsilon}$ are estimates obtained from the above U.S. mortality rates. The empirical sizes ($d = 0$) and empirical powers ($d = -1, -2, -3, -4, -5$) of the proposed test Z are computed based on 10,000 replications, and are reported in Tables 4–6.

Table 4 shows that (i) the size becomes more accurate when T is larger, (ii) it is a bit larger than the nominal level for a smaller T and (iii) the size for Group II is more accurate than that for Group I since age groups at the two ends are believed to be quite volatile. Tables 5 and 6 show that the proposed test has non-trivial powers, and for a larger T, the power increases as $|d|$ becomes bigger. Note that we do not expect the power will increase as T becomes large for a fixed

TABLE 4

EMPIRICAL SIZE. WE COMPUTE THE EMPIRICAL SIZE OF THE PROPOSED TEST Z AT NOMINAL LEVEL $\alpha = 0.05, 0.1$ BASED ON 10, 000 REPETITIONS FOR SAMPLE SIZE $T = 50, 100, 200, 500$ FROM (2) AND (3) WITH $\phi_1 = 1$ AND $\phi_2 = 0$.

	$\alpha = 0.05$				$\alpha = 0.1$			
	$T = 50$	$T = 100$	$T = 200$	$T = 500$	$T = 50$	$T = 100$	$T = 200$	$T = 500$
Group I	0.1483	0.0951	0.0692	0.0579	0.2280	0.1583	0.1311	0.1135
Group II	0.0912	0.0625	0.0578	0.0532	0.1453	0.1155	0.1114	0.1062

TABLE 5

EMPIRICAL POWER FOR GROUP I. WE REPORT THE EMPIRICAL POWER OF THE PROPOSED TEST Z AT NOMINAL LEVEL $\alpha = 0.05, 0.1$ BASED ON 10, 000 REPETITIONS FOR SAMPLE SIZE $T = 50, 100, 200, 500$ FROM (2) AND (3) WITH $\phi_1 = 1 + d/\sqrt{T}$ & $\phi_1 + \phi_2 = 1 + d/T^{3/2}$.

	$\alpha = 0.05$				$\alpha = 0.1$			
d	$T = 50$	$T = 100$	$T = 200$	$T = 500$	$T = 50$	$T = 100$	$T = 200$	$T = 500$
−1	0.2960	0.1980	0.1300	0.1017	0.4074	0.2853	0.2215	0.1753
−2	0.4804	0.3385	0.2478	0.1931	0.5846	0.4472	0.3513	0.2966
−3	0.6384	0.5146	0.3962	0.3220	0.7187	0.6206	0.5124	0.4310
−4	0.7113	0.6868	0.5534	0.4740	0.7515	0.7665	0.6589	0.5908
−5	0.6801	0.7600	0.7235	0.6122	0.6940	0.8030	0.8079	0.7107

TABLE 6

EMPIRICAL POWER FOR GROUP II. WE REPORT THE EMPIRICAL POWER OF THE PROPOSED TEST Z AT NOMINAL LEVEL $\alpha = 0.05, 0.1$ BASED ON 10, 000 REPETITIONS FOR SAMPLE SIZE $T = 50, 100, 200, 500$ FROM (2) AND (3) WITH $\phi_1 = 1 + d/\sqrt{T}$ & $\phi_1 + \phi_2 = 1 + d/T^{3/2}$.

	$\alpha = 0.05$				$\alpha = 0.1$			
d	$T = 50$	$T = 100$	$T = 200$	$T = 500$	$T = 50$	$T = 100$	$T = 200$	$T = 500$
−1	0.1517	0.1146	0.0989	0.0956	0.2205	0.1892	0.1675	0.1606
−2	0.2866	0.2236	0.2089	0.2000	0.3727	0.3226	0.3092	0.2983
−3	0.4570	0.4008	0.3709	0.3644	0.5378	0.5065	0.4758	0.4793
−4	0.5242	0.6047	0.5541	0.5437	0.5802	0.6961	0.6598	0.6631
−5	0.5018	0.7494	0.7397	0.7170	0.5379	0.8025	0.8118	0.8077

d since the alternative hypothesis depends on both d and T. Since the proposed test requires a large T to ensure a reasonable size and the mortality index plays a key role in forecasting mortality risk, it is important to develop a more powerful test for a unit root in the Lee–Carter model in the future.

4. Conclusions

Although the Lee–Carter mortality model and its extensions have been extensively applied in demography and actuarial science, asymptotic properties of its statistical inference remain unknown. Recently, Leng and Peng (2016) proved that the two-step inference procedure proposed by Lee and Carter (1992) may lead to a wrong identification of the dynamics of the mortality index when the mortality index does not follow from a unit root AR(1) process. A blind application of the two-step inference procedure to the U.S. mortality rate data leads to the conclusion that the mortality index follows from a unit root AR(1) model.

By assuming that the mortality index follows from an AR(p) model, this paper proposes a method to test whether it is a unit root AR(1) model. Due to the special structure of errors, asymptotic distribution derived for the proposed test is quite different from existing ones in the literature of econometrics. An application of this test to the U.S. mortality rate data rejects the null hypothesis that the mortality index follows from a unit root AR(1) model, which contradicts the blind application of the two-step inference procedure proposed by Lee and Carter (1992). Therefore, the proposed test is useful in checking the applicability of the widely employed Lee–Carter mortality model and its extensions. On the other hand, the simulation study indicates that the proposed test requires a large number of observations ($\geq$100) to ensure a reasonable size, it becomes important to develop a more powerful test for a unit root in the Lee–Carter model in the future.

Acknowledgements

We thank a reviewer for his/her useful comments. Peng's research was partly supported by the Simons Foundation and a CAE grant from SOA.

References

BAUER, D., BENTH, F.E. and KIESEL, R. (2012) Modeling the forward surface of mortality. *SIAM Journal on Financial Mathematics,* **3**, 639–666.

BISETTI, E. and FAVERO, C.A. (2014) Measuring the impact of longevity risk on pension systems: The case of Italy. *North American Actuarial Journal,* **18**, 87–103.

BROUHNS, N., DENUIT, M. and VERMUNT, J.K. (2002) A Poisson log-linear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics,* **31**, 373–393.

CAIRNS, A.J., BLAKE, D. and DOWD, K. (2008) Modelling and management of mortality risk: A review. *Scandinavian Actuarial Journal,* **2008**, 79–113.

CAIRNS, A.J., BLAKE, D., DOWD, K., COUGHLAN, G.D., EPSTEIN, D. and KHALAF-ALLAH, M. (2011) Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics,* **48**, 355–367.

CHEN, H. and COX, S.H. (2009) Modeling mortality with jumps: Applications to mortality securitization. *Journal of Risk and Insurance,* **76**, 727–751.

CURRIE, I.D., DURBAN, M. and EILERS, P.H.C. (2004) Smoothing and forecasting mortality rates. *Statistical Modelling,* **4**, 279–298.

D'AMATO, V., HABERMAN, S., PISCOPO, G., RUSSOLILLO, M. and TRAPANI, L. (2014) Detecting common longevity trends by a multiple population approach. *North American Actuarial Journal,* **18**, 139–149.

FREES, E.W., CARRIERE, J.F. and VALDEZ, E.A. (1996) Annuity valuation with dependent mortality. *Journal of Risk and Insurance,* **63**, 229–261.

GIROSI, F. and KING, G. (2007) *Understanding the Lee–Carter Mortality Forecasting Method.* Working Paper, Harvard University. Available at https://gking.harvard.edu/files/abs/lc-abs.shtml.

HABERMAN, S. and RENSHAW, A.E. (2008) Mortality, longevity and experiments with the Lee–Carter model. *Lifetime Data Analysis,* **14**, 286–315.

KÖLBL, F. (2006) Aggregation of AR(2) processes. *Technical report at Technische Universität Graz.* Available at www.stat.tugraz.at/dthesis/koelbl06.pdf.

LEE, R. and CARTER, L. (1992) Modelling and forecasting US mortality. *Journal of the American Statistical Association,* **87**, 659–671.

LENG, X. and PENG, L. (2016) Inference pitfalls in Lee–Carter model for forecasting mortality. *Insurance: Mathematics and Economics,* **70**, 58–65.

LI, N. and LEE, R. (2005) Coherent mortality forecasts for a group of populations: An extension of the Lee–Carter method. *Demography,* **42**, 575–594.

LIN, Y., TAN, K.S., TIAN, R. and YU, J. (2014) Downside risk management of a defined benefit plan considering longevity basis risk. *North American Actuarial Journal,* **18**, 68–86.

NJENGA, C.N. and SHERRIS, M. (2011) Longevity risk and the econometric analysis of mortality trends and volatility. *Asia-Pacific Journal of Risk and Insurance,* **5**, 24–76.

PHILLIPS, P.C.B. and DURLAUF, S.N. (1986) Multiple time series regression with integrated processes. *The Review of Economic Studies,* **53**, 473–495.

PHILLIPS, P.C.B. (1987) Towards a unified asymptotic theory for autoregression. *Biometrika,* **74**, 535–547.

PHILLIPS, P.C.B. and PERRON, P. (1988) Testing for a unit root in time series regression. *Biometrika,* **75**, 335–346.

SAID, S.E. and DICKEY, D. (1984) Testing for unit roots in autoregressive moving-average models with unknown order. *Biometrika,* **71**, 599–607.

XIAO, Z. (2014) Unit roots: A selective review of the contributions of Peter C.B. Phillips. *Econometric Theory,* **30**, 775–814.

YANG, S.S. and WANG, C.W. (2013) Pricing and securitization of multi-country longevity risk with mortality dependence. *Insurance: Mathematics and Economics,* **52**, 157–169.

ZHANG, L., SHEN, H. and HUANG, J. (2013) Robust regularized singular value decomposition with application to mortality data. *Annals of Applied Statistics,* **7**, 1540–1561.

XUAN LENG
Erasmus University Rotterdam
P.O. Box 1738
3000 DR Rotterdam, The Netherlands
E-Mail: leng@ese.eur.nl

LIANG PENG (Corresponding author)
Department of Risk Management and Insurance
Robinson College of Business
Georgia State University
Atlanta, GA 30303, USA
E-Mail: lpeng@gsu.edu

APPENDIX A

A.1. Proof of Theorem 1

Before we prove the main theorem, we need some lemmas.

Lemma 1. *Suppose conditions of Theorem 1 hold.*

i. *If $\delta^\tau\delta = 0$, then for $\phi_1 = 1 + d_1/\sqrt{T}$ & $\phi_1 + \phi_2 = 1 + d_2/T$ and any fixed i, we have*

ia. $T^{-1/2} y^{(i)}_{[sT]} \xrightarrow{D} \Sigma^{1/2}\tilde{J}_{d_2}(s)$ *in space* $D([0, 1])$;

ib. $T^{-2}\sum_{t=1}^{T} y_t^{(i)} y_t^{(i)\tau} \xrightarrow{d} \Sigma^{1/2}\int_0^1 \tilde{J}_{d_2}(s)\tilde{J}^\tau_{d_2}(s)\,ds\,\Sigma^{1/2}$;

ic. $T^{-5/2}\sum_{t=1}^{T} t y_t^{(i)} \xrightarrow{d} \Sigma^{1/2}\int_0^1 s\tilde{J}_{d_2}(s)\,ds$;

id.

$$T^{-1}\sum_{t=1}^{T}\left(y_t^{(i+1)}u^{*\tau}_{t-i} + u^*_{t-i}y_t^{(i+1)\tau}\right)$$
$$\xrightarrow{d} \Sigma^{1/2}\left(\tilde{J}_{d_2}(1)\tilde{J}^\tau_{d_2}(1) - \tilde{J}_{d_2}(0)\tilde{J}^\tau_{d_2}(0) - 2(d_2 + d_1^2)\int_0^1 \tilde{J}_{d_2}(s)\tilde{J}^\tau_{d_2}(s)\,ds\right)\Sigma^{1/2} - \Sigma - 2\Sigma^\epsilon.$$

ii. *If $\delta^\tau\delta > 0$, then for $\phi_1 = 1 + d_1/\sqrt{T}$ & $\phi_1 + \phi_2 = 1 + d_2/T^{3/2}$ and any fixed i, we have*

iia. $T^{-3}\sum_{t=1}^{T} y_t^{(i)} y_t^{(i)\tau} \xrightarrow{p} \frac{1}{12}\delta\delta^\tau$;

iib. $T^{-3}\sum_{t=1}^{T} t y_t^{(i)} \xrightarrow{p} \frac{1}{12}\delta$;

iic. $T^{-3/2}\sum_{t=1}^{T} u_t^* y_t^{(i)\tau} \xrightarrow{d} \Sigma^{1/2}(\frac{W(1)}{2} - \int_0^1 W(s)\,ds)\delta^\tau$.

Proof 1. Write

$$\begin{pmatrix}\log m_{x,t}\\ \log m_{x,t-1}\end{pmatrix} = \begin{pmatrix}\phi_1 & \phi_2\\ 1 & 0\end{pmatrix}\begin{pmatrix}\log m_{x,t-1}\\ \log m_{x,t-2}\end{pmatrix} + \begin{pmatrix}\delta_x + u_{x,t}\\ 0\end{pmatrix},$$

and it follows from iterations that

$$\begin{pmatrix}\log m_{x,t}\\ \log m_{x,t-1}\end{pmatrix} = \sum_{i=1}^{t}\begin{pmatrix}\phi_1 & \phi_2\\ 1 & 0\end{pmatrix}^{t-i}\begin{pmatrix}\delta_x + u_{x,i}\\ 0\end{pmatrix}.$$

Using the arguments in Kölbl (2006), we have

$$\log m_{x,t} = \sum_{i=1}^{t-1}\frac{\lambda_1^{i+1} - \lambda_2^{i+1}}{\lambda_1 - \lambda_2}(\delta_x + u_{x,t-i}), \tag{A1}$$

where

$$\lambda_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} \quad\text{and}\quad \lambda_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}.$$

First, we consider the case of $\delta^\tau\delta = 0$. In this case,

$$\lambda_1 = 1 + \frac{d_2}{T} + o(T^{-1}) \quad\text{and}\quad \lambda_2 = \frac{d_1}{\sqrt{T}} + o(T^{-1/2}). \tag{A2}$$

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

It follows from (A1), (A2), (5) and the same arguments in Phillips (1987) that

$$\frac{\Sigma^{-1/2}}{\sqrt{T}}\log m_{[sT]}\xrightarrow{D}\int_0^s e^{d_2(s-t)}\,dW(t)=J_{d_2}(s),$$

i.e.,

$$\frac{\Sigma^{-1/2}}{\sqrt{T}}y_{[sT]}^{(0)}\xrightarrow{D}J_{d_2}(s)-\int_0^1 J_{d_2}(t)\,dt=\tilde{J}_{d_2}(s). \tag{A3}$$

Hence, (ia)–(ic) follows from (A3) easily. For proving (id), it follows from (4) that

$$y_t^{(0)}=\phi_1 y_t^{(1)}+\phi_2 y_t^{(2)}+u_t^*. \tag{A4}$$

So, we have

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^T y_t^{(0)}y_t^{(0)\tau}\\
&=\frac{1}{T}\sum_{t=1}^T\left(\phi_1 y_t^{(1)}+\phi_2 y_t^{(2)}+u_t^*\right)\left(\phi_1 y_t^{(1)}+\phi_2 y_t^{(2)}+u_t^*\right)^\tau\\
&=\frac{1}{T}\sum_{t=1}^T\phi_1^2 y_t^{(1)}y_t^{(1)\tau}+\frac{1}{T}\sum_{t=1}^T\phi_2^2 y_t^{(2)}y_t^{(2)\tau}+\frac{1}{T}\sum_{t=1}^T u_t^*u_t^{*\tau}\\
&\quad+\frac{1}{T}\sum_{t=1}^T\phi_1\phi_2\left(y_t^{(1)}y_t^{(2)\tau}+y_t^{(2)}y_t^{(1)\tau}\right)+\frac{1}{T}\sum_{t=1}^T\phi_1\left(y_t^{(1)}u_t^{*\tau}+u_t^*y_t^{(1)\tau}\right)\\
&\quad+\frac{1}{T}\sum_{t=1}^T\phi_2\left(y_t^{(2)}u_t^{*\tau}+u_t^*y_t^{(2)\tau}\right)\\
&=\frac{1}{T}\sum_{t=1}^T\left(1+\frac{2d_1}{\sqrt{T}}+\frac{d_1^2}{T}\right)y_t^{(1)}y_t^{(1)\tau}+\frac{1}{T}\sum_{t=1}^T\frac{d_1^2}{T}y_t^{(2)}y_t^{(2)\tau}+\frac{1}{T}\sum_{t=1}^T u_t^*u_t^{*\tau}\\
&\quad+\frac{1}{T}\sum_{t=1}^T\left(\frac{d_2}{T}-\frac{d_1}{\sqrt{T}}-\frac{d_1^2}{T}\right)\left(y_t^{(1)}y_t^{(2)\tau}+y_t^{(2)}y_t^{(1)\tau}\right)\\
&\quad+\frac{1}{T}\sum_{t=1}^T\left(y_t^{(1)}u_t^{*\tau}+u_t^*y_t^{(1)\tau}\right)\left(1+o_p(1)\right)+o_p(1)\\
&=\frac{1}{T}\sum_{t=1}^T y_t^{(1)}y_t^{(1)\tau}+\frac{1}{T}\sum_{t=1}^T\left(\frac{2d_1}{\sqrt{T}}+\frac{d_1^2}{T}\right)\left(1+\frac{d_1}{\sqrt{T}}\right)^2 y_t^{(2)}y_t^{(2)\tau}\\
&\quad+\frac{1}{T}\sum_{t=1}^T\frac{d_1^2}{T}y_t^{(2)}y_t^{(2)\tau}+\frac{1}{T}\sum_{t=1}^T u_t^*u_t^{*\tau}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{T}\sum_{t=1}^{T}\left(\frac{d_2}{T}-\frac{d_1}{\sqrt{T}}-\frac{d_1^2}{T}\right)\left(1+\frac{d_1}{\sqrt{T}}\right)\left(y_t^{(2)}y_t^{(2)\tau}+y_t^{(2)}y_t^{(2)\tau}\right)\\
&+\frac{1}{T}\sum_{t=1}^{T}\left(y_t^{(1)}u_t^{*\tau}+u_t^{*}y_t^{(1)\tau}\right)\left(1+o_p(1)\right)+o_p(1)\\
&=\frac{1}{T}\sum_{t=1}^{T}y_t^{(1)}y_t^{(1)\tau}+\frac{1}{T}\sum_{t=1}^{T}\frac{2(d_1^2+d_2)}{T}y_t^{(2)}y_t^{(2)\tau}\\
&+\frac{1}{T}\sum_{t=1}^{T}u_t^{*}u_t^{*\tau}+\frac{1}{T}\sum_{t=1}^{T}\left(y_t^{(1)}u_t^{*\tau}+u_t^{*}y_t^{(1)\tau}\right)\left(1+o_p(1)\right)+o_p(1),
\end{aligned}
$$

which implies (id) with $i=0$ by using (ia), (ib) and $T^{-1}\sum_{t=1}^{T}u_t^{*}u_t^{*\tau}\xrightarrow{p}\Sigma+2\Sigma^{\epsilon}$. Similarly, we can show (id) holds for $i\geq 1$.

Next, we consider the case of $\delta^{\tau}\delta>0$. In this case,

$$
\lambda_1=1+\frac{d_2}{T^{3/2}}+o\left(T^{-3/2}\right)\quad\text{and}\quad\lambda_2=\frac{d_1}{\sqrt{T}}+o\left(T^{-1/2}\right). \tag{A5}
$$

Then, it follows from (A1) and (A5) that

$$
T^{-1}\log m_{[sT]}\xrightarrow{p}s\delta,
$$

i.e.,

$$
T^{-1}y_{[sT]}^{(0)}\xrightarrow{p}\left(s-\int_0^1 y\,dy\right)\delta=(s-1/2)\delta. \tag{A6}
$$

Hence, we can show (iia)–(iic) by using (A6). ■

Proof of Theorem 1 2. Define

$$
\begin{aligned}
A_{i,j}&:=\sum_{x=1}^{K}\sum_{t=1}^{T}\left(y_{x,t-i}^{(0)}-\phi_1 y_{x,t-i}^{(1)}-\phi_2 y_{x,t-i}^{(2)}\right)\left(y_{x,t-j}^{(0)}-\phi_1 y_{x,t-j}^{(1)}-\phi_2 y_{x,t-j}^{(2)}\right)\\
&=\sum_{x=1}^{K}\sum_{t=1}^{T}u_{x,t-i}^{*}u_{x,t-j}^{*}
\end{aligned}
$$

and

$$
B_{i,j}:=\sum_{x=1}^{K}\sum_{t=1}^{T}y_{x,t}^{(j)}\left(y_{x,t-i}^{(0)}-\phi_1 y_{x,t-i}^{(1)}-\phi_2 y_{x,t-i}^{(2)}\right)=\sum_{x=1}^{K}\sum_{t=1}^{T}u_{x,t-i}^{*}y_{x,t}^{(j)}.
$$

i. It follows from the weak law of large numbers that, as $T\to\infty$,

$$
\begin{aligned}
T^{-1}A_{i,j}&=\sum_{x=1}^{K}T^{-1}\sum_{t=1}^{T}u_{x,t-i}^{*}u_{x,t-j}^{*}\\
&\xrightarrow{p}\sum_{x=1}^{K}\mathsf{E}\left(\beta_x e_{t-i}+\epsilon_{x,t-i}-\epsilon_{x,t-i-1}\right)\left(\beta_x e_{t-j}+\epsilon_{x,t-j}-\epsilon_{x,t-j-1}\right)\\
&=\begin{cases}\sum_{x=1}^{K}(\sigma_{x,x}+2\sigma_{x,x}^{\epsilon}) & \text{if}\quad i=j\\ -\sum_{x=1}^{K}\sigma_{x,x}^{\epsilon} & \text{if}\quad |i-j|=1\\ 0 & \text{if}\quad |i-j|>1,\end{cases}
\end{aligned} \tag{A7}
$$

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

$$\lim_{T\to\infty} T^{-1}\mathsf{E}\left\{\sum_{x=1}^{K}\sum_{t=1}^{T} u^{*}_{x,t}u^{*}_{x,t-2}\right\}^2$$

$$= \lim_{T\to\infty} T^{-1}\mathsf{E}\left\{\sum_{x=1}^{K}\sum_{t=1}^{T} u_{x,t}u_{x,t-2}\right\}^2$$

$$= \lim_{T\to\infty}\sum_{x_1=1}^{K}\sum_{x_2=1}^{K}\frac{1}{T}\sum_{t=1}^{T}\left\{\mathsf{E}\left(u_{x_1,t}u_{x_2,t}\right)\right\}^2$$

$$+ \lim_{T\to\infty}\sum_{x_1=1}^{K}\sum_{x_2=1}^{K}\frac{2}{T}\sum_{t=1}^{T}\mathsf{E}\left(u_{x_1,t}u_{x_1,t-2}u_{x_2,t-1}u_{x_2,t-3}\right)$$

$$= \sum_{x_1=1}^{K}\sum_{x_2=1}^{K}\left\{(\sigma_{x_1,x_2}+2\sigma^{\epsilon}_{x_1,x_2})^2+2(\sigma^{\epsilon}_{x_1,x_2})^2\right\}.$$

and

$$\frac{1}{\sqrt{T}}A_{0,2}\xrightarrow{d} N\left(0,\sum_{x_1=1}^{K}\sum_{x_2=1}^{K}\left\{(\sigma_{x_1,x_2}+2\sigma^{\epsilon}_{x_1,x_2})^2+2(\sigma^{\epsilon}_{x_1,x_2})^2\right\}\right), \tag{A8}$$

which is independent of $W(s)$. It follows from Lemma 1 (i) and (A7) that

$$T^{-2}D_{i,i} = \mathrm{tr}\left(T^{-2}\sum_{t=1}^{T} y_t^{(i)}y_t^{(i)\tau}\right)\xrightarrow{d}\mathrm{tr}\left(\Sigma\int_0^1 \tilde{J}_{d_2}(s)\tilde{J}^{\tau}_{d_2}(s)\,ds\right), \tag{A9}$$

$$\begin{aligned}
T^{-1}B_{i,i+1} &= \mathrm{tr}\left(T^{-1}\sum_{t=1}^{T} u^{*}_{t-i}y_t^{(i+1)\tau}\right)\\
&= \frac{1}{2}\mathrm{tr}\left(T^{-1}\sum_{t=1}^{T}\left(u^{*}_{t-i}y_t^{(i+1)\tau}+y_t^{(i+1)}u^{*\tau}_{t-i}\right)\right)\\
&\xrightarrow{d}\frac{1}{2}\mathrm{tr}\left(\Sigma\left(\tilde{J}_{d_2}(1)\tilde{J}^{\tau}_{d_2}(1)-\tilde{J}_{d_2}(0)\tilde{J}^{\tau}_{d_2}(0)-2(d_2+d_1^2)\int_0^1\tilde{J}_{d_2}(s)\tilde{J}^{\tau}_{d_2}(s)\,ds\right)-\Sigma-2\Sigma^{\epsilon}\right)\\
&= Z_2,
\end{aligned} \tag{A10}$$

$$\begin{aligned}
T^{-1}B_{i,i} &= T^{-1}\sum_{x=1}^{K}\sum_{t=1}^{T}u^{*}_{x,t-i}\left(y^{(i)}_{x,t}-y^{(i+1)}_{x,t}\right)+T^{-1}B_{i,i+1}\\
&\xrightarrow{d}\sum_{x=1}^{K}(\sigma_{x,x}+2\sigma^{\epsilon}_{x,x})+Z_2
\end{aligned} \tag{A11}$$

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

and

$$T^{-1}B_{i,i+2} = T^{-1}\sum_{x=1}^{K}\sum_{t=1}^{T}u^{*}_{x,t-i}y^{(i+1)}_{x,t} - T^{-1}\sum_{x=1}^{K}\sum_{t=1}^{T}u^{*}_{x,t-i}\left(y^{(i+1)}_{x,t} - y^{(i+2)}_{x,t}\right)$$

$$\xrightarrow{d} Z_2 + \sum_{x=1}^{K}\sigma^{\epsilon}_{xx}. \tag{A12}$$

Now, using (A4), (A9), (A10), (A7) and Lemma 1 (i), we have

$$\begin{aligned} D_{0,1} - D_{1,1} &= \sum_{x=1}^{K}\sum_{t=1}^{T}y^{(1)}_{x,t}\left(y^{(0)}_{x,t} - y^{(1)}_{x,t}\right) \\ &= \sum_{x=1}^{K}\sum_{t=1}^{T}u^{*}_{x,t}y^{(1)}_{x,t} + \sum_{x=1}^{K}\sum_{t=1}^{T}y^{(1)}_{x,t}\left((\phi_1 - 1)y^{(1)}_{x,t} + \phi_2 y^{(2)}_{x,t}\right) \\ &= B_{0,1} + (\phi_1 - 1)\sum_{x=1}^{K}\sum_{t=1}^{T}y^{(1)}_{x,t}\left(y^{(1)}_{x,t} - y^{(2)}_{x,t}\right) + (\phi_1 + \phi_2 - 1)D_{1,2} \\ &= O_p(T). \end{aligned} \tag{A13}$$

Similarly, we have

$$\begin{aligned} D_{1,1} - D_{1,2} &= B_{1,1} + (\phi_1 - 1)\sum_{x=1}^{K}\sum_{t=1}^{T}y^{(1)}_{x,t}\left(y^{(2)}_{x,t} - y^{(3)}_{x,t}\right) \\ &\quad + (\phi_1 + \phi_2 - 1)D_{1,3} + o_p(1) \\ &= O_p(T), \end{aligned} \tag{A14}$$

$$\begin{aligned} D_{1,2} - D_{2,2} &= B_{1,2} + (\phi_1 - 1)\sum_{x=1}^{K}\sum_{t=1}^{T}y^{(2)}_{x,t}\left(y^{(2)}_{x,t} - y^{(3)}_{x,t}\right) \\ &\quad + (\phi_1 + \phi_2 - 1)D_{2,3} + o_p(1) \\ &= O_p(T) \end{aligned} \tag{A15}$$

and

$$\begin{aligned} D_{0,2} - D_{1,2} &= B_{0,2} + (\phi_1 - 1)\sum_{x=1}^{K}\sum_{t=1}^{T}y^{(2)}_{x,t}\left(y^{(1)}_{x,t} - y^{(2)}_{x,t}\right) \\ &\quad + (\phi_1 + \phi_2 - 1)D_{2,2} + o_p(1) \\ &= O_p(T), \end{aligned} \tag{A16}$$

which imply that

$$D_{1,1} - D_{1,2} - (D_{1,2} - D_{2,2}) = B_{1,1} - B_{1,2} + o_p(T) = A_{1,1} + o_p(T).$$

Therefore, we obtain

$$\begin{aligned} & D_{1,1} D_{2,2} - D_{1,2}^2 \\ &= D_{1,2}[D_{1,1} - D_{1,2} - (D_{1,2} - D_{2,2})] - (D_{1,1} - D_{1,2})(D_{1,2} - D_{2,2}) \\ &= D_{1,2} A_{1,1} + o_p(T^3), \end{aligned} \tag{A17}$$

$$\begin{aligned} & D_{0,1} D_{2,2} - D_{0,2} D_{1,2} - D_{12} \sum_{x=1}^{K} \sum_{t=1}^{T} \left(y_{x,t}^{(0)} - y_{x,t}^{(1)}\right) \left(y_{x,t}^{(1)} - y_{x,t-1}^{(2)}\right) - (D_{1,1} D_{2,2} - D_{1,2}^2) \\ &= (D_{0,1} - D_{1,1})(D_{2,2} - D_{1,2}) - D_{1,2} \sum_{x=1}^{K} \sum_{t=1}^{T} \left(y_{x,t}^{(0)} - y_{x,t}^{(1)}\right) \left(y_{x,t}^{(2)} - y_{x,t-1}^{(2)}\right) \\ &= O_p(T^2) - D_{12} \Bigg[\sum_{x=1}^{K} \sum_{t=1}^{T} u_{x,t}^{*} \left(y_{x,t}^{(2)} - y_{x,t-1}^{(2)}\right) \\ &\quad + \frac{d_1}{\sqrt{T}} \sum_{x=1}^{K} \sum_{t=1}^{T} \left(y_{x,t}^{(1)} - y_{x,t}^{(2)}\right) \left(y_{x,t}^{(2)} - y_{x,t-1}^{(2)}\right) + o_p(\sqrt{T}) \Bigg] \\ &= -D_{1,2} \left(A_{0,2} + \frac{d_1}{\sqrt{T}} A_{1,2} \right) + o_p(T^{5/2}), \end{aligned} \tag{A18}$$

and

$$\begin{aligned} & D_{0,1} D_{2,2} - D_{0,2} D_{1,2} + D_{1,1} D_{0,2} - D_{0,1} D_{1,2} - (D_{1,1} D_{2,2} - D_{1,2}^2) \\ &= -(D_{0,1} - D_{1,1})(D_{1,2} - D_{2,2}) + (D_{1,1} - D_{1,2})(D_{0,2} - D_{1,2}) \\ &= -\left(B_{0,1} + \frac{d_2}{T} D_{1,2} \right) \left(B_{1,2} + \frac{d_2}{T} D_{2,3} \right) + \left(B_{1,1} + \frac{d_2}{T} D_{1,3} \right) \left(B_{0,2} + \frac{d_2}{T} D_{2,2} \right) + o_p(T^2) \\ &= -\left(B_{0,1} + \frac{d_2}{T} D_{1,2} \right) \left(B_{1,2} + \frac{d_2}{T} D_{1,2} \right) + \left(B_{1,1} + \frac{d_2}{T} D_{1,2} \right) \left(B_{0,2} + \frac{d_2}{T} D_{1,2} \right) + o_p(T^2) \\ &= -B_{0,1} B_{1,2} + B_{1,1} B_{0,2} + \frac{d_2}{T} (-B_{0,1} - B_{1,2} + B_{1,1} + B_{0,2}) D_{1,2} + o_p(T^2) \\ &= -A_{0,1} B_{1,2} + A_{1,1} B_{0,2} + \frac{d_2}{T} D_{1,2} (A_{1,1} - A_{0,1}) + o_p(T^2). \end{aligned} \tag{A19}$$

Hence, Theorem 1 (i) follows from (A17)–(A19).

ii. It follows from Lemma 1 (ii) that

$$T^{-3} D_{i,j} \xrightarrow{p} \frac{\delta^{\tau} \delta}{12} \tag{A20}$$

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press

and

$$T^{-3/2} B_{i,j} \xrightarrow{d} \mathrm{tr}\left(\Sigma^{1/2} \left(\frac{W(1)}{2} - \int_0^1 W(s)\, ds \right) \delta^\tau \right). \tag{A21}$$

As before, using (A20) and (A21), we can show that

$$D_{1,1} D_{2,2} - D_{1,2}^2 = D_{1,2} A_{1,1} + o_p(T^4),$$

$$D_{0,1} D_{2,2} - D_{0,2} D_{1,2} - D_{1,2} \sum_{x=1}^{K} \sum_{t=1}^{T} \left(y_{x,t}^{(0)} - y_{x,t}^{(1)} \right) \left(y_{x,t}^{(1)} - y_{x,t-1}^{(2)} \right) - (D_{1,1} D_{2,2} - D_{1,2}^2)$$
$$= \left(A_{0,2} + \tfrac{d_1}{\sqrt{T}} A_{1,2} \right) D_{1,2} + o_p(T^{7/2}),$$

and

$$D_{0,1} D_{2,2} - D_{0,2} D_{1,2} + D_{1,1} D_{0,2} - D_{0,1} D_{1,2} - (D_{1,1} D_{2,2} - D_{1,2}^2)$$

$$= -A_{0,1} B_{1,2} + A_{1,1} B_{0,2} + \frac{d_2}{T^{3/2}} (A_{1,1} - A_{0,1}) D_{1,2} + o_p(T^{5/2}),$$

which imply Theorem 1 (ii) by noting that

$$\mathsf{E} \left(\frac{W_i(1)}{2} - \int_0^1 W_i(s)\, ds \right)^2 = \frac{1}{4} - \int_0^1 s\, ds + \int_0^1 \int_0^1 \min(s, t)\, ds dt = \frac{1}{4} - \frac{1}{2} + \frac{1}{3} = \frac{1}{12}$$

and

$$\mathsf{E}\, Z_4^2 = \mathsf{E} \left\{ \mathrm{tr} \left(\delta^\tau \Sigma^{1/2} \left(\frac{W(1)}{2} - \int_0^1 W(s)\, ds \right) \left(\frac{W^\tau(1)}{2} - \int_0^1 W^\tau(s)\, ds \right) \Sigma^{1/2} \delta \right) \right\}$$

$$= \mathrm{tr} \left(\mathsf{E} \left[\left(\frac{W(1)}{2} - \int_0^1 W(s)\, ds \right) \left(\frac{W^\tau(1)}{2} - \int_0^1 W^\tau(s)\, ds \right) \right] \Sigma^{1/2} \delta \delta^\tau \Sigma^{1/2} \right)$$

$$= \frac{1}{12} \mathrm{tr} \left(I_{K \times K} \Sigma^{1/2} \delta \delta^\tau \Sigma^{1/2} \right)$$

$$= \frac{1}{12} \mathrm{tr} \left(\Sigma \delta \delta^\tau \right).$$

■

https://doi.org/10.1017/asb.2017.24 Published online by Cambridge University Press