TI 2001-078/4
Tinbergen Institute Discussion Paper

Block Local to Unity and Continuous Record Asymptotics

H. Peter Boswijk

Department of Quantitative Economics, Faculty of Economics and Econometrics, University of Amsterdam

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Block local to unity and continuous record asymptotics

H. Peter Boswijk
Department of Quantitative Economics, Universiteit van Amsterdam*

August 2001

Abstract

This paper provides a continuous record interpretation of the block local to unity asymptotics proposed recently by Phillips, Moon and Xiao (2001). It also demonstrates that in the case of homogeneous dynamics and a fixed number of blocks, the new asymptotic approximation coincides with the conventional local to unity asymptotic approximation.

1 Introduction

Consider the first-order autoregressive model for an observed time series $\{y_t, t = 0, 1, \ldots, n\}$:

$$y_t = ay_{t-1} + u_t, \qquad t = 1, \ldots, n, \tag{1}$$

where u_t is an i.i.d. $(0, \sigma^2)$ process. It is well known that the asymptotic properties of the least-squares estimator $\hat{a} = \left(\sum_{t=1}^n y_{t-1}^2\right)^{-1} \sum_{t=1}^n y_{t-1}y_t$ are quite different for the three cases $|a| < 1$, $a = 1$ and $a > 1$, corresponding to stationary, integrated and explosive processes, respectively, in terms of both the speed of convergence and the shape of the limiting distribution of the normalized estimator. This difference, and in particular the discontinuity of the asymptotic properties at $a = 1$, has motivated the development of *local to unity* asymptotics, where y_t is regarded as a triangular array generated by (1) with

$$a = e^{c/n} \approx 1 + \frac{c}{n}, \tag{2}$$

with $c \in \mathbb{R}$ constant for all n, and hence $a \to 1$ as $n \to \infty$. This was first proposed by Bobkoski (1983), and subsequently Phillips (1987a) and Chan and Wei (1987) showed that this leads to asymptotic behaviour of $\hat{a}$ that changes continuously from $c < 0$ (hence $a < 1$) via $c = 0$ ($a = 1$) to $c > 0$ ($a > 1$). Furthermore, Phillips (1987a) showed that the conventional (fixed a) asymptotics are recovered from this approach by letting $c \to \pm\infty$.

In a recent paper, Phillips, Moon and Xiao (2001), henceforth PMX, propose a generalization of the local to unity approach, which they term *block local to unity* asymptotics. The basic idea is to think of

*Address for correspondence: Department of Quantitative Economics, Universiteit van Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands. E-mail: peterb@fee.uva.nl.

the sample of size n as consisting of M blocks of m observations each, such that $n = Mm$, and to replace (2) by

$$a = e^{c/m} \approx 1 + \frac{c}{m}. \tag{3}$$

Hence $a \to 1$ as the number of observations m within a block approaches infinity. PMX show that the asymptotic distribution of $\hat{a}$ for M fixed and $m \to \infty$ is a generalization of the local-to-unity asymptotic distribution, characterized by a finite diffusion mixture. On the other hand, under *sequential asymptotics* where $m \to \infty$ first, followed by $M \to \infty$ (denoted by $(m, M \to \infty)_{seq}$), it is shown that when $c < 0$, $m(\hat{a}-1)$ is a $\sqrt{M}$-consistent and asymptotically normal estimator of c. This ability to deliver consistent estimation of c is viewed as one of the main advantages of the block local to unity approach, because the lack of convergence of $n(\hat{a} - 1)$ to c as $n \to \infty$ under (2) leads to serious problems for bootstrapping (see Basawa *et al.*, 1991) as well as for inference on cointegration with near-integrated processes (see Elliott, 1998). An additional advantage is that the approach allows for a continuum of convergence rates of $\hat{a}$ under stationarity ($c < 0$): letting $m = n^{\gamma}$ and $M = n^{1-\gamma}$ with $0 \leq \gamma \leq 1$, it follows that $\hat{a} - a = \mathrm{O}_p\left(M^{-1/2}m^{-1}\right) = \mathrm{O}_p\left(n^{-1/2-\gamma/2}\right)$, which varies from $\mathrm{O}_p\left(n^{-1/2}\right)$ ($\gamma = 0$, fixed a asymptotics) through $\mathrm{O}_p(n^{-1})$ ($\gamma = 1$, local to unity asymptotics). PMX also analyse extensions of the basic model, allowing for various starting value assumptions, for short-run dependence in u_t with possibly changing parameters over the blocks, and for deterministic regressors in (1).

In this paper, I provide an interpretation of the block local to unity parametrization from a continuous record asymptotic perspective, see Phillips (1987a, b), Perron (1991a, b) and Sørensen (1992). In particular, it is shown that y_t generated by (1) under (3) arises quite naturally as a discretely sampled continuous-time Ornstein-Uhlenbeck process on $[0, M]$. On the one hand, this gives an additional justification for (3), but at the same time it indicates that the $M \to \infty$ assumption may not yield a very accurate approximation in some practical situations. For fixed M and homogeneous dynamics, I will demonstrate that inference based on the block local to unity approach essentially coincides with the conventional local to unity inference, although the former involves a more natural parametrization.

The plan of the remainder of this paper is as follows. Section 2 sets out the continuous-time framework and shows how this provides a convenient interpretation of the block local to unity model under i.i.d. innovations. Section 3 discusses how the results change in the presence of heterogeneous dynamics, and the final section discusses the practical relevance of the results. The notation used in the paper is largely the same as in PMX, with the exception of "$\sim$" denoting "is distributed as"; weak convergence is denoted by "$\Rightarrow$".

2 Continuous record asymptotics

Consider a bivariate Brownian motion $(B(s), B_{-1}(s))'$ on $[0, \infty)$ with variance matrix $\sigma^2 I_2$, and define the Ornstein-Uhlenbeck (OU) process $J_c(s)$ on $[0, \infty)$, generated conditional on its starting value $J_c(0)$ by

$$dJ_c(s) = cJ_c(s)ds + dB(s). \tag{4}$$

The starting value $J_c(0)$ is determined by $B_{-1}(s), s \in [0,\infty)$, independently of $B(s), s \in [0,\infty)$, by either one of the two following possibilities:

(i) $J_c(0) = \int_{-\kappa}^{0} e^{-(s+\kappa)c} dB_{-1}(-s) \sim N(0, \sigma^2(e^{2\kappa c} - 1)/2c)$ for some finite $\kappa \geq 0$.

(ii) $c < 0$ and $J_c(0) = \int_{-\infty}^{0} e^{-sc} dB_{-1}(-s) \sim N(0, -\sigma^2/2c)$.

Assumption (i) corresponds to the process starting at zero at time $s = -\kappa$; possibility (ii) states that the process started at the infinite past, and hence has the stationary $N(0, -\sigma^2/2c)$ distribution at time $s = 0$.

In what follows it will be helpful to think of the time index s as calendar time, e.g. measured in years. Suppose that we observe J_c over the period $[0, M]$, at $n+1$ equidistant time points

$$s_0 = 0, \quad s_1 = \frac{1}{n}M, \quad \ldots, \quad s_{n-1} = \frac{n-1}{n}M, \quad s_n = M.$$

Suppose also that we choose $n = mM$ for some integer m, such that $s_t = t/m, t = 0, \ldots, n$. Then m is the number of time points in an interval of unit length, and hence, if s is measured in years, the number of observations per year. A well known property of the OU process is

$$\begin{aligned} J_c(s_t) &= e^{cs_t} J_c(0) + \int_0^{s_t} e^{(s_t - s)c} dB(s) \\ &= e^{c/m} \left(e^{cs_{t-1}} J_c(0) + \int_0^{s_{t-1}} e^{(s_{t-1} - s)c} dB(s) \right) + \int_{s_{t-1}}^{s_t} e^{(s_t - s)c} dB(s) \\ &= e^{c/m} J_c(s_{t-1}) + \int_{s_{t-1}}^{s_t} e^{(s_t - s)c} dB(s). \end{aligned} \tag{5}$$

The final term is independent of $J_c(s_{t-1})$ and has a $N(0, \sigma^2(e^{2c/m} - 1)/2c)$ distribution when $c \neq 0$, and a $N(0, \sigma^2/m)$ distribution for $c = 0$. Thus, let

$$y_t = \tau_{c,m} J_c(s_t), \qquad u_t = \tau_{c,m} \int_{s_{t-1}}^{s_t} e^{(s_t - s)c} dB(s), \qquad t = 1, \ldots, n, \tag{6}$$

with $y_0 = \tau_{c,m} J_c(0)$, and where

$$\tau_{c,m} = \begin{cases} \sqrt{2c/(e^{2c/m} - 1)}, & c \neq 0, \\ \sqrt{m}, & c = 0. \end{cases}$$

Then is it clear that y_t satisfies the first-order autoregression (1) with $a = e^{c/m}$ and $u_t \sim$ i.i.d. $N(0, \sigma^2)$. Note that the scaling factor $\tau_{c,m}$ is used only to obtain correspondence between the variances of u_t and $B(s)$; it could be replaced simply by $\sqrt{m}$ without affecting any of the results below, since the least-squares estimator is invariant to a scale change in $\{y_t\}$.

The above discussion shows that the block local to unity model, in its simplest form, arises naturally when the time series $\{y_t\}$ is a discretely observed OU process over a period $[0, M]$. As an example of where such a situation might occur, suppose that we have observations on a short-term interest rate over a period of M years, at various frequencies, e.g. monthly ($m = 12$), weekly ($m = 52$) or daily ($m = 250$, allowing for weekends and bank holidays). The well-known Vasicek (1977) model entails

that the short rate follows an OU process, and this model is often used to price interest rate derivatives, which requires an estimate of c (and σ). Taking the obvious estimator $\hat{c} = m \log \hat{a} \approx m(\hat{a} - 1)$ of c, one would like to assess the estimation uncertainty of $\hat{c}$, and in particular the effect of the observation frequency m and the time span M on this estimation error. The block local to unity approach allows us to study these effects separately, and shows that for accurate estimation of c it is important to choose the time span M sufficiently large, since $(\hat{c} - c) = \mathrm{O}_p(M^{-1/2})$. (Note that consistent estimation of σ requires only $m \to \infty$.) This confirms the empirical experience that the mean-reversion in interest rates is typically so weak that, e.g., ten years of data is not sufficient to obtain a reliable estimate of c (significantly different from 0), even if one has high-frequency data and hence a very large sample size n.

Consider now the asymptotic behaviour of $\hat{a}$ when y_t is generated by (6). When m is fixed and $M \to \infty$, then this is simply a fixed-parameter autoregression, and hence the classical asymptotic properties for the cases $a < 1$, $a = 1$ and $a > 1$ emerge (for the explosive case, special care has to be taken of the starting value assumption). Focussing on the stationary case, this means that as $M \to \infty$ (fixed m),

$$\sqrt{n}(\hat{a} - a) \Rightarrow N(0, 1 - a^2), \tag{7}$$

which implies

$$\sqrt{M}(\hat{c} - c) = \sqrt{m}\sqrt{n}(\hat{a} - a) + \mathrm{o}_p(1) \Rightarrow N(0, m(1 - e^{2c/m})). \tag{8}$$

Note that when $m \to \infty$ after $M \to \infty$, this implies $\sqrt{M}(\hat{c} - c) \Rightarrow N(0, -2c)$, which is identical to the $(m, M \to \infty)_{seq}$ result obtained by PMX (even though the limit operations are interchanged here). Note also that the role of m is relatively minor in this result, as the difference between $m(1 - e^{2c/m})$ and $-2c$ will be relatively small for realistic values of c and m.

Quite different results are obtained for M fixed, $m \to \infty$. In that case it follows for all c that

$$m(\hat{a} - a) = \frac{\frac{1}{\tau_{c,m}^2} \sum_{t=1}^{n} y_{t-1} u_t}{\frac{1}{m\tau_{c,m}^2} \sum_{t=1}^{n} y_{t-1}^2} = \frac{\sum_{t=1}^{n} J_c(s_{t-1}) \int_{s_{t-1}}^{s_t} e^{(s_t - s)c} dB(s)}{\frac{1}{m} \sum_{t=1}^{n} J_c(s_{t-1})^2} \Rightarrow \frac{\int_0^M J_c(s) dB(s)}{\int_0^M J_c(s)^2 ds}. \tag{9}$$

The main step in proving this limit result is the fact that the difference between $\int_{s_{t-1}}^{s_t} e^{(s_t - s)c} dB(s)$ and $\int_{s_{t-1}}^{s_t} dB(s) = B(s_t) - B(s_{t-1})$ vanishes as $m \to \infty$, which is easily checked using integration by parts on $\int_{s_{t-1}}^{s_t} e^{(s_t - s)c} dB(s)$. The limiting expression (9) is very resemblant to the usual local to unity limiting distribution of $n(\hat{a} - a)$, the main difference being that the integrals are over $[0, M]$ instead of the unit interval. However, a simple time scale change reveals that the two approaches are actually equivalent: letting $r = s/M$, it follows that $\bar{B}(r) = M^{-1/2} B(rM)$ is a Brownian motion on $[0, 1]$ with variance σ^2, and $\bar{J}_{\bar{c}}(r) = M^{-1/2} J_c(rM)$ is an OU process on $[0, 1]$ with mean-reversion parameter $\bar{c} = Mc$, generated by

$$d\bar{J}_{\bar{c}}(r) = \bar{c} \bar{J}_{\bar{c}}(r) dr + d\bar{B}(r),$$

with $\bar{J}_{\bar{c}}(0) = M^{-1/2}J_c(0)$. Therefore, (9) may be rewritten as

$$n(\hat{a}-a) = Mm(\hat{a}-a) \Rightarrow \frac{\int_0^M J_c(s)dB(s)}{\frac{1}{M}\int_0^M J_c(s)^2 ds} = \frac{\int_0^1 \bar{J}_{\bar{c}}(r)d\bar{B}(r)}{\int_0^1 \bar{J}_{\bar{c}}(r)^2 dr}. \tag{10}$$

This shows that the same asymptotic inference applies as in the usual local to unity case, using an appropriate definition of $\bar{c} = n\log a = Mm\log a = Mc$, and an appropriate starting value assumption, namely that the process has started at zero at time $-\bar{\kappa} = -\kappa/M$, with $\bar{\kappa}$ either finite (case (i)) or infinite with $c < 0$ (case (ii)). The asymptotic distribution (10) is invariant to σ but does depend on the starting value (even when $c = 0$), in particular on the choice of $\bar{\kappa}$, but this is a well-known property of continuous record asymptotics, see, e.g., Phillips (1987b) and Perron (1991b).

The asymptotic normality under $(m, M \to \infty)_{seq}$ with $c < 0$ now follows directly from (9), since the right-hand side expression equals $\hat{c}_{ML} - c$, the centered maximum likelihood estimator of c based on a continuous sample $\{J_c(s), s \in [0, M]\}$. When $c < 0$, this estimator is well known to be $\sqrt{M}$-consistent and asymptotically normal, with $\sqrt{M}(\hat{c}_{ML} - c) \Rightarrow N(0, -2c)$, see e.g. Basawa and Prakasa Rao (1980). As shown by PMX, this result essentially follows from writing the integrals from 0 to M as the sum of integrals from 0 to 1, and applying a central limit theorem to the numerator of (10) (divided by $\sqrt{M}$) and a law of large numbers to the denominator. Note that, using (10) and $\bar{c} = Mc$, this actually provides an alternative proof of Phillips' (1987a) result that as $\bar{c} \to -\infty$,

$$\sqrt{\frac{-1}{2\bar{c}}}\frac{\int_0^1 \bar{J}_{\bar{c}}(r)d\bar{B}(r)}{\int_0^1 \bar{J}_{\bar{c}}(r)^2 dr} \Rightarrow N(0,1).$$

The above discussions have shown that this sequential asymptotic result coincides with the result from applying $M \to \infty$ first, followed by $m \to \infty$, and furthermore it has appeared that the continuous record step ($m \to \infty$) has a minor role in this, and only slightly changes the variance of the asymptotic normal distribution.

In summary, it may be concluded that in the case of i.i.d. innovations, the new block local to unity asymptotic approach coincides with either the standard local to unity approach (if the number of blocks M is kept fixed), or with the traditional fixed a asymptotics (under $(m, M \to \infty)_{seq}$), at least in the mean-reverting ($c < 0$) and unit root ($c = 0$) case. However, PMX also consider the case where u_t is generated by a possibly heterogeneous linear process, the implications of which are discussed in the next section.

3 Heterogeneous dynamics

PMX assume that instead of an i.i.d. $N(0, \sigma^2)$ process, u_t is modelled as a heterogeneous linear process $u_t = \sum_{j=0}^{\infty} b_{k,j}\varepsilon_{t-j}$, where $\varepsilon_t \sim$ i.i.d. $(0, 1)$ and $0 < (\sum_{j=0}^{\infty} b_{k,j})^2 = \omega_k^2 < \infty$, and where $k \in \{0, 1, \ldots, M\}$ is a block index. This implies that y_t may no longer be interpreted as a discretely sampled

OU process, but the same continuous record notion may still be used to interpret the block local to unity asymptotic results. The serial dependence in u_t implies that the least-squares estimator has an asymptotic bias term, but this may be corrected in the usual non-parametric way as shown by PMX. More important is the heterogeneity in the long-run variances ω_k^2, which implies that $m^{-1/2}y_{[sm]}$ no longer converges weakly to a homogeneous OU process, but instead converges to a process $H_c(s)$ generated by

$$dH_c(s) = cH_c(s)ds + \omega(s)dW(s), \tag{11}$$

where $W(s)$ is a standard Brownian motion on $[0, M]$ and $\omega(s)^2 = \sum_{k=1}^{M} 1_{(k-1,k]}(s)\omega_k^2$. The process is started up in the same way as J_c (possibly also allowing for some pre-sampe heterogeneity). Note that PMX do not use the notation in terms of one process on $[0, M]$, but write $H_c(s)$ as a concatenated version of the OU processes $H_{k,c}(r)$ on $[0, 1]$, each driven by its own Brownian motion $B_k(r)$ with variance ω_k^2. The M fixed, $m \to \infty$ limit theory for the least-squares estimator corrected for bias, denoted $\hat{a}^+$, now may be expressed as

$$m(\hat{a}^+ - a) \Rightarrow \frac{\int_0^M H_c(s)\omega(s)dW(s)}{\int_0^M H_c(s)^2 ds}. \tag{12}$$

This limiting distribution now depends not only on c but also on $\{\omega_k^2\}$, even as $M \to \infty$, see Theorem 2 of PMX. An obvious improvement over the ordinary least-squares estimator is the (feasible) weighted least-squares estimator $\hat{a}_\omega^+$, which has the following limit distribution as $m \to \infty$:

$$m(\hat{a}_\omega^+ - a) \Rightarrow \frac{\int_0^M \omega(s)^{-1}H_c(s)dW(s)}{\int_0^M \omega(s)^{-2}H_c(s)^2 ds}. \tag{13}$$

This distribution still depends on $\{\omega_k^2\}$, because $\omega(s)^{-1}H_c(s)$ is not a homogeneous OU process; note that H_c is standardized by its most recent volatility, but determined itself by the current and all previous volatilities. Because $\{\omega_k^2\}$ can be estimated consistently ($m \to \infty$), this need not pose an insurmountable problem for inference on a, as the distribution in (13) can be obtained given c and $\{\omega_k^2\}$ by simulation. However, it indicates a lack of robustness of the conventional local to unity asymptotics to this type of heterogeneity. If $c < 0$ and the average variance $M^{-1}\sum_{k=1}^{M}\omega_k^2$ converges to a constant, then as $M \to \infty$, this problem disappears, since $\sqrt{M}m(\hat{a}_\omega^+ - a) \Rightarrow N(0, -2c)$ just as in the homogeneous case, see PMX, Theorem 3.

4 Discussion

The purpose of this paper has been to provide a reinterpretation of the block local to unity asymptotics in a continuous record framework, which allows the number of blocks M to be interpreted as the time span, and m as the observation frequency. The block local to unity approach allows us to study the

effect of these two parameters on the distribution of the least-squares estimator separately, and one might argue that this is the main contribution of this approach. The continuum of convergence rates n^{α}, $\alpha \in [\frac{1}{2}, 1]$ under stationarity can also be interpreted in the continuous record framework: the rate of convergence is determined by the rate at which the information $\sigma^{-2}\sum_{t=1}^{n} y_{t-1}^2$ grows, and the continuous record asymptotic analysis shows that (if $c < 0$) this quantity grows linearly with the time span M but quadratically with the sampling frequency m. Thus the rate at which the information grows with the sample size depends on the relative contribution of sampling frequency and time span to the growth of n. Put differently, the relevant result is $\hat{a} - a = \mathrm{O}_p\left(M^{-1/2}m^{-1}\right)$, which is more informative than $\hat{a} - a = \mathrm{O}_p(n^{-\alpha})$ for some α.

The continuous record interpretation has shown that under homogeneity, the block local to unity approach either coincides with the local to unity approach (fixed time span) or yields virtually the same inference as traditional fixed a asymptotics, at least for $a \leq 1$. Under block heterogeneity the model does yield new results, and in particular highlights the lack of robustness of local to unity inference to this type of heterogeneity. Although one might argue that the model where the variance is constant over time stretches of equal length (say, years) is not very realistic for economic data, the result is indicative of the type of problems encountered with local to unity asymptotics when the volatility is time-varying but displays some persistence, such that the heteroskedasticity is not averaged out unless one considers a very long time span. A very related result was obtained by Boswijk (2000) for the case where the volatility follows a near-integrated GARCH processes. The present approach shows that in such cases one either has to use the estimated volatility process for obtaining critical values, or revert to the large M asymptotic normal approximation (assuming $a < 1$).

The question that remains is which type of asymptotics practitioners should be advised to use. In the interest rate example mentioned earlier, it seems likely that 10 years of daily data is not sufficient to use the asymptotic normal approximation for the distribution of $\hat{c}$. For macro-economic time series such as the real GDP one might have reliable data over a considerably longer time span, but then the true rate of mean reversion might be so close to zero that the normal approximation is again not reliable. From (10) we see that it is $\bar{c} = Mc$ rather than M only that determines the accuracy of the normal approximation, which again brings us back to the original local to unity result. This also casts some doubt on the ability of this new approach to solve the problems with bootstrapping and cointegration inference mentioned in the introduction, since these will be primarily dependent on $\bar{c}$ rather than c. It is clear that an asymptotic normal confidence interval $\hat{c} \pm 1.96\sqrt{-2\hat{c}/M}$ is invalid if it contains $c = 0$, since it would then contain at least one parameter value which is accepted based on the wrong null distribution. But how far the upper bound of the interval should be away from zero to yield reliable inferences based on asymptotic normality is an open question, determined by the question how close to the normal (in some suitable sense) the distribution of (10) is. Since the local to unity inference will automatically converge to normal inference as $Mc \to -\infty$, the most reliable way to proceed in empirical practice would be to use the former at all times, taking appropriate care of possible dependence and heterogeneity.

References

Basawa, I.V., A.K. Lallik, W.P. McCormick, J.H. Reeves and R.L. Taylor (1991), "Bootstrapping unstable first order autoregressive process", *Annals of Statistics*, 19, 1098–1101.

Basawa, I.V. and B.L.S. Prakasa Rao (1980), *Statistical Inference for Stochastic Processes*. London: Academic Press.

Bobkoski, M.J. (1983), *Hypothesis Testing in Nonstationary Time Series*. Ph.D. Thesis, University of Wisconsin.

Boswijk, H.P. (2000), "Testing for a unit root with near-integrated volatility", Working paper, Universiteit van Amsterdam, `http://www.fee.uva.nl/ke/boswijk/urnivol.pdf`.

Chan, N.H. and C.Z. Wei (1987), "Asymptotic inference for nearly nonstationary AR(1) processes". *Annals of Statistics*, 15, 1050–1063.

Elliott, G. (1998), "On the robustness of cointegration methods when regressors almost have unit roots", *Econometrica*, 66, 149–158.

Perron, P. (1991a), "A continuous time approximation to the unstable first-order autoregressive model: the case without an intercept", *Econometrica*, 59, 211-236.

Perron, P. (1991b), "A continuous time approximation to the stationary first-order autoregressive model", *Econometric Theory*, 7, 236–252.

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

Phillips, P.C.B. (1987b), "Time series regression with a unit root", *Econometrica*, 55, 277–301.

Phillips, P.C.B., H.R. Moon and Z. Xiao (2001), "How to estimate autoregressive roots near unity", *Econometric Theory*, 17, 29–69.

Sørensen, B.E. (1992), "Continuous record asymptotics in systems of stochastic differential equations", *Econometric Theory*, 8, 28–51.

Vasicek, O. (1977), "An equilibrium characterization of the term structure", *Journal of Financial Economics*, 5, 177–188.