

UDC 519.21

J. RODERICK MCCRORIE

A ROLE OF THE SKOROKHOD SPACE IN THE DEVELOPMENT OF THE ECONOMETRIC ANALYSIS OF TIME SERIES

This paper discusses the fundamental role played by the Skorokhod space, through its underpinning of functional central limit theory, in the development of the paradigm of unit roots and co-integration. This paradigm has fundamentally affected the way economists approach economic time series as was recognized by the award of the Nobel Memorial Prize in Economic Sciences to Robert F. Engle and Clive W.J. Granger in 2003. Here, we focus on how P.C.B. Phillips and others used the Skorokhod topology to establish a limiting distribution theory that underpinned and facilitated the development of methods of estimation and testing of single equations and systems of equations with possibly integrated regressors. This approach has spawned a large body of work that can be traced back to Skorokhod's conception of fifty years ago. Much of this work is surprisingly confined to the econometrics literature.

1. INTRODUCTION

One important aspect of a time series is that it is but one realization of a multidimensional random variable. From this point of view, an assumption of second-order stationarity is convenient because it facilitates the inference through laws of large numbers and central limit theorems in a classical way. An early influence on models of economic time series was rendered by the book by Grenander and Rosenblatt [35] which was based on the stationarity about deterministic trend functions, where the inference can be conducted under what today would be called “Grenander conditions” (see, e.g., [43, p. 215]). These include a requirement that a suitably normalized sample moment matrix of regressors converge to a positive definite limit. Another aspect of the second-order stationarity assumption is that it permits a Wold decomposition whereby a time series can be represented as the sum of a linearly regular part involving an infinite-order weighted average of white noise and a part that is perfectly linearly deterministic [43, p. 137]. This offered some justification for the then emerging Box–Jenkins methods [9], where autoregressive integrated moving average (ARIMA) models selected on the basis of the data could be viewed as approximations to the regular part in the Wold decomposition. These models produced satisfactory representations of many observed economic time series, at least for the purpose of prediction. But, as the models were selected purely on the basis of the data, they lacked a theoretical justification, as if they emerged from a “black box”. The essential problem faced by the econometrics profession in the late 1960s and 1970s was that structural econometric models, embodying restrictions from economic theory, were often outperformed by the black box models. For example, Cooper [14] compared one-step-ahead forecasts from seven structural models of the U.S. economy against a naïve forecast from an AR model and found forecasts from the latter were superior in

2000 *AMS Mathematics Subject Classification*. Primary 60F17, 60G50, 60J15, 60J55, 60J65, 62M10, 62P20.

Key words and phrases. Skorokhod space, functional central limit theorems, non-stationary time series, unit roots and co-integration, Wiener functionals, econometrics.

most cases. Naylor et al. [69] even advocated Box–Jenkins models as an alternative to econometric models. There was also the problem of the proliferation of data-based models for a given time series that were incompatible with one another from the point of view of economic theory. In the background, there was a debate about whether or not the observed aggregate time series, which were manifestly non-stationary, should have random walk or other trend components removed prior to the estimation given the loss of a “long run” information that such transformations imply. The introduction of the concepts of unit roots and co-integration in the context of non-stationary time series [37,30,38,39,44] helped to resolve some of these issues. The theoretical underpinning of this work was provided in papers by P.C.B. Phillips [76,75,83] who applied functional central limit theory based on the Skorokhod space in a way that has firmly established it as a part of the econometrician’s toolkit. The result has been a distinct literature that traces back to Skorokhod’s conception of fifty years ago. The purpose of this paper is to summarize its main contributions for the wider mathematicians’ community.

In Section 2, we briefly discuss the concepts and antecedents of unit roots and co-integration, and error correction models which Engle and Granger [30] associated with co-integration. In Section 3, we show how Phillips used the Skorokhod topology to underpin the time series regression with a unit root and how he provided an (asymptotic) explanation of influential simulation results on spurious regression by Granger and Newbold [40]. Regression with co-integrated variables is then briefly contrasted with spurious regression. In Section 4, we discuss the recent work, focusing on functional central limit theory appropriate for non-linear transformations of non-stationary time series. To our knowledge, this is the first article to describe this literature as a whole. For reasons of space, we confine ourselves almost exclusively to a discussion of autoregressive unit roots. This remains the most important application in econometrics (e.g., [41]), but other work is relevant, especially under fractional integration [17,19,6,52,61,62,97–99,60,34].

2. UNIT ROOTS AND CO-INTEGRATION

Intuitively, co-integration tries to bring sense to the notion that two or more time series when looked at individually appear to be wandering erratically and yet are systematically wandering together. Casual inspection of many raw economic time series would suggest they exhibit a trend component, explained possibly by the time series being stationary about a deterministic trend or possibly because they are non-stationary, say as a result of their containing of a random walk component. The following idea is difficult to make precise ([18,57,94,10,32,33,69]), but we could imagine, analogously to the Box–Jenkins methodology, of “differencing” a time series to induce “stationarity”. More formally, we could say that if, asymptotically, a time series comprises a sequence of random variables with the first and second moments which tend to fixed stationary values and whose covariances tend to stationary values depending only on how far apart the elements are, then the time series is integrated to order zero, or $I(0)$. A time series could then be said to be integrated to order d , or $I(d)$, if it must be differenced d times before an $I(0)$ series results. This would mean, then, that the random walk process

$$(1) \quad x_t = x_{t-1} + \varepsilon_t, \quad x_0 = 0, \quad \varepsilon_t \sim IID(0, \sigma^2)$$

is such that it is $I(1)$, because the series of its differences, $\nabla x_t = x_t - x_{t-1}$, is $I(0)$. See [45,46] for a non-technical discussion and the comparison of the properties of $I(1)$ and $I(0)$ series. Much of the early literature was concerned with formulating and interpreting the tests to discriminate between stationary processes about a deterministic trend and non-stationary processes containing an autoregressive unit root such as (1) above [28]. This was motivated as a means of trying to avoid applying a wrong filter (either one of detrending or differencing) to the data which, it was argued, would result in the

invalid inference. It is, however, becoming more common today to see both types of the model as simply alternative representations of the components of an underlying, trending stochastic process [81,63].

The idea of co-integration, when applied in the context of I(1) and I(0) series, seeks to relate raw series that are I(1) through a linear combination of them which is I(0). We could imagine series that individually exhibit “persistence” but which together are attracted to each other towards a statistical equilibrium where linear combinations of the variables are “stationary”. Co-integration does not explain why the variables are trending in the first place, but what it does do is to provide a basis of examining the relationships between I(1) variables or variables integrated to higher order. Often we can place an economic interpretation on the co-integrating relation. Engle and Granger [30,38] showed that every co-integrated system can be written as a model that has what is called an error correction representation, comprising a “balanced” equation in I(0) variables involving the differences of the variables related to their lagged values but also to a term that reflects the disequilibrium between the I(1) variables which, under co-integration, is I(0). Co-integration is a property that can be applied, in principle, to time series with possibly higher orders of integration, and various representations of co-integrated systems are useful [86,108,79,42,111].

3. REGRESSION WITH NON-STATIONARY TIME SERIES

The essential statistical distributional problem in unit root regression, which Chan and Wei [11] and Phillips [77] related to the convergence in the Skorokhod topology, can be exposted in the context of the simple AR(1) model

$$(2) \quad x_t = \rho x_{t-1} + \varepsilon_t \quad (t = 1, \dots, n),$$

where ρ is the unknown parameter, and $\varepsilon_t \sim NID(0, 1)$. Consider the joint density

$$(3) \quad f_\rho(x) = (\sqrt{2\pi})^{-n} \exp \left\{ -\frac{1}{2} [(1 + \rho)^2 T_2 - 2\rho T_1 + x_n^2] \right\},$$

where

$$(4) \quad T_1 = \sum_{t=1}^n x_t x_{t-1}, \quad T_2 = \sum_{t=1}^n x_{t-1}^2.$$

The maximum-likelihood estimator (MLE) $\hat{\rho}_n$ of ρ is given by

$$(5) \quad \hat{\rho}_n = T_1/T_2.$$

If $|\rho| < 1$, the process is stationary, and [59]

$$(6) \quad \sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{d} N(0, 1 - \rho^2),$$

where \xrightarrow{d} denotes the convergence in distribution. When $\rho = 1$, (6) is not useful as a basis of testing for a unit root against stationarity; however, it can be shown using a standard argument (e.g., [42]) that

$$(7) \quad n^{-1} \sum_{t=1}^n x_{t-1} \varepsilon_t \xrightarrow{d} \frac{1}{2}(\chi_1^2 - 1),$$

which with

$$(8) \quad \sum_{t=1}^n x_{t-1}^2 = O(n^2)$$

suggests writing

$$(9) \quad n(\hat{\rho}_n - 1) = n^{-1} \sum_{t=1}^n x_{t-1} \varepsilon_t / n^{-2} \sum_{t=1}^n x_{t-1}^2,$$

on substituting (2) into (3). The problem that presents itself is that there is no law of large numbers with this normalization such that the denominator converges to a constant. What (9) does convey is that if the statistics has a well-defined limit distribution, which Theorem 1A below confirms, the ML estimator converges, in this case, to this distribution at a faster rate than it does to a normal distribution in the stationary case. The asymptotic sampling properties of the MLE in (5) can be described in the following way, which is more suggestive of a uniform approach.

Theorem 1A. ([59, 112, 3, 24]) *Let $x_0 = 0$. Then*

$$(10) \quad \sqrt{I_n(\rho)}(\hat{\rho}_n - \rho) \xrightarrow{d} \begin{cases} N, & |\rho| < 1, \\ \rho \frac{W^2(1) - 1}{2^{3/2} \int_0^1 W^2(r) dr}, & |\rho| = 1, \\ C, & |\rho| > 1, \end{cases}$$

where N is a centered Gaussian random variable with variance 1, C is a standard Cauchy random variable, $W = (W(r), r \geq 0)$ is a standard Wiener process, and $I_n(\rho)$ is the (expected) Fisher information contained in x_1, \dots, x_n about the parameter ρ , given by $E_\rho(T_2)$. As $n \rightarrow \infty$,

$$(11) \quad E_\rho(T_2) \sim \begin{cases} \frac{n}{1 - \rho^2}, & |\rho| < 1, \\ \frac{n^2}{2}, & |\rho| = 1, \\ \frac{\rho^{2n}}{(\rho^2 - 1)^2}, & |\rho| > 1. \end{cases}$$

If $|\rho| \leq 1$, (10) remains valid under the weaker assumptions that x_0 is an arbitrary constant or a random variable with a finite second moment not depending on the sequence $\varepsilon = (\varepsilon_t, t \geq 1)$, and ε a arbitrary sequence of centered and normalized independent, identically distributed (i.i.d.) random variables. If $|\rho| > 1$, the limit distribution depends on the initial value and, in general, on the particular distribution of each ε_t , even if ε forms a sequence of i.i.d. random variables [56]. Different results hold if ε forms a sequence of i.i.d. random variables with a stable distribution or in the domain of attraction of a stable law [12,78,67].

If we normalize by the observed rather than the expected Fisher information, the three limit distributions are reduced to two:

Theorem 1B. ([59, 3, 24]) *Let $x_0 = 0$. Then*

$$(12) \quad \sqrt{I_n(\rho)}(\hat{\rho}_n - \rho) \xrightarrow{d} \begin{cases} N, & |\rho| < 1, \\ \rho \frac{W^2(1) - 1}{2\sqrt{\int_0^1 W^2(r) dr}}, & |\rho| = 1. \end{cases}$$

The essential contribution of Chan and Wei [11] and Phillips [77] in the journal paper immediately following [30] was to set the convergence results in the case $|\rho| = 1$ to the

context of functional central limit theory. Phillips's approach, which was more obviously designed as an extension of classical asymptotic theory, has led to the now prevailing two-stage approach towards deriving the limit distributions pertaining to unit root statistics. This involves using, in the first stage, the functional central limit theorem (FCLT) to derive a limit distribution for the (normalized) integrated process itself and then, in the second stage, to derive the limit distribution of the sample statistics based explicitly on its construction as a functional of the integrated process. It is in the first stage that the ideas of Skorokhod are applied because it is here that the classical theory, pertaining to the convergence in a distribution of random variables, needs to be extended to a framework involving the weak convergence of random functions. This is needed because the persistence in random shocks requires us to consider the *whole trajectory* of the process and not just its endpoint.

Consider a real-valued stochastic process $(x_t, t \in \mathcal{N})$. We should like to consider the weak convergence in a function space that is complete, to avoid the probability mass escaping from the space as $n \rightarrow \infty$, and separable, because then all the Borel sets of the space are measurable, and the weak convergence on product spaces is equivalent to the weak convergence on the component spaces. It is possible to use the space of continuous functions on the unit interval, $C[0, 1]$, endowed with the uniform metric but, because most of the functions of interest are not continuous, this involves an awkward construction that requires extra terms that are defined to make the relevant partial sum process continuous and are shown to be asymptotically negligible [17,15]. Phillips [77,76,84] based his work, instead, on the space $D[0, 1]$ of cadlag (*continue à droite, limites à gauche*) functions, which contains jumps but not isolated points, but is sufficient for the problem in hand. This space is not separable under the uniform metric, meaning in practice there are "too many" sets, on which a probability space can be defined. The problem is therefore precisely the one Skorokhod [105] resolved fifty years ago: the space can be rendered separable by what is now popularly called the Skorokhod metric (his J_1 metric) defined in such a way that allows functions to be compared "sideways," as well as vertically: for $x, y \in D[0, 1]$,

$$(13) \quad d_S = \inf_{\lambda \in \Lambda} \left\{ \varepsilon > 0 : \sup_t |\lambda(t) - t| \leq \varepsilon, \sup_t |x(t) - y(\lambda(t))| \leq \varepsilon \right\},$$

where Λ denotes the set of functions $\lambda : [0, 1] \rightarrow [0, 1]$. This is the crucial element in terms of what is needed to define the Borel σ -algebra; however, the metric space $(D[0, 1], d_S)$ is not complete, and Phillips used a modification of (13) introduced by Billingsley (1968) that preserves the same topology:

$$(14) \quad d_B = \inf_{\lambda \in \Lambda'} \left\{ \varepsilon > 0 : \|\lambda\| \leq \varepsilon, \sup_t |x(t) - y(\lambda(t))| \leq \varepsilon \right\},$$

where

$$(15) \quad \|\lambda\| = \sup_{t \neq s} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|,$$

and Λ' denotes the set of all increasing functions such that $\|\lambda\| \leq \infty$.

The two-stage approach to establishing the limiting distributions of sample statistics considers, in the first stage, a real-valued stochastic process $(x_t, t \in \mathcal{N})$ such that $n^{-1/2}x_{[rn]} \Rightarrow \sigma W(r)$, $\sigma > 0$, where $[rn]$, $r \in [0, 1]$, denotes the integer part of rn , " \Rightarrow " denotes the weak convergence in $D[0, 1]$, as described above, and W represents the standard Brownian motion on $[0, 1]$. The approach is set up such that one of the variety of such FCLT's could be employed [26,31,65,66,48–50,113]. With econometric time series in mind, Phillips chose to use the conditions by Herrndorff [49] which allow for certain

types of weakly dependent and heterogeneously distributed disturbances. Later, Phillips and Solo [91] advocated an approach based on ε following a linear process. Today, we might instead use the conditions based on results by Doukhan and Louhichi [21,4], Beare [5], de Jong and Davidson [22], or Davidson [16].

In the second stage, an argument based on the continuous mapping theorem [95] is used to show that, for continuous real-valued functions T on \mathcal{R} ,

$$(16) \quad n^{-1} \sum_{t=1}^n T(n^{-1/2}x_t) \xrightarrow{d} \int_0^1 T(\sigma W(r))dr.$$

For the denominator in (8),

$$(17) \quad \begin{aligned} n^{-2} \sum_{t=1}^n x_{t-1}^2 &= n^{-1} \sum_{t=1}^n \left(n^{-1/2} \sum_{s=1}^{t-1} \varepsilon_s \right)^2 \\ &= n^{-1} \sum_{t=1}^n \left(n^{-1/2} \sum_{s=1}^{[rn]} \varepsilon_s \right)^2 \quad \left(\frac{t-1}{n} \leq r \leq \frac{t}{n} \right) \\ &\xrightarrow{d} \int_0^1 W(r)^2 dr, \end{aligned}$$

using (16). For the numerator,

$$(18) \quad \begin{aligned} n^{-1} \sum_{t=1}^n x_{t-1} \varepsilon_t &= \sum_{t=1}^n \left(n^{-1/2} \sum_{s=1}^{t-1} \varepsilon_s \right) n^{-1/2} \varepsilon_t \\ &\xrightarrow{d} \sigma^2 \int_0^1 W(r) dW(r). \end{aligned}$$

The integral in (18) is an Itô integral equal to $\frac{1}{2}(W^2(1) - 1)$, which agrees with (7). The joint convergence of (17) and (18) gives (10) up to a normalization.

Expressions involving functionals of Brownian motion such as (10) and (12) can be derived using the above two-stage approach for many relevant time series models, the most important being the unit root model that includes a constant, and a constant and a time trend (see [15, 42 (Ch. 17)]). With the densities of (10) and (12), we could use, in principle, a transformation theorem *to generate* the densities pertaining to the other models [1]. Deriving and computing these densities, however, is not a tractable problem either analytically or numerically. Indeed, except in the simple model above [2,64], we do not have analytic expressions for the asymptotic moments, densities, or distributions. This means, in practice, critical values for hypothesis tests are constructed by simulating the densities from the Wiener functionals, and, for general ARMA models, additional methods need to be employed, often using a decomposition by Beveridge and Nelson [7] under linear process assumptions [91]. This gives what are called Phillips–Perron tests [90] and augmented Dickey–Fuller tests (see [42, Ch. 17]), [107]). More recently, in an attempt to improve the asymptotic nature of expressions such as (10) and (12), computationally intensive methods based on the bootstrap [73,72] have been designed.

Perhaps the most celebrated application of functional limit theory based on the weak convergence in the Skorokhod topology to econometrics was the explanation offered by Phillips [76] of simulation results by Granger and Newbold [40]. Specifically, suppose we estimate the parameters α and β by OLS in the model

$$(19) \quad y_t = \alpha + \beta x_t + \varepsilon_t \quad (t = 1, \dots, n)$$

where both x_t and y_t are generated by independent random walks. Granger and Newbold found in experiments on a sample sizes of 50 and 100 (constrained by the then limits of computer technology) that the coefficient of determination, R^2 , was very high (in the standard setting, indicating that most of the variation in y is explained by the variation in x), the Durbin–Watson (DW) statistics was extremely low (in the standard setting, suggesting among other possibilities the residuals were serially correlated), and tests of the significance of β were seriously biased towards the *rejection of no relationship* between y and x . Phillips showed under the FCLT that

$$(20) \quad \hat{\beta} \xrightarrow{d} \left(\frac{\sigma_v}{\sigma_w} \right) \frac{S_{XY}}{S_{XX}}, \quad n^{-1/2}t_\beta \xrightarrow{d} \frac{S_{XY}}{\sqrt{S_{XX}S_{YY} - S_{XY}^2}},$$

$$(21) \quad R^2 \xrightarrow{d} \frac{S_{XY}^2}{S_{XX}S_{YY}}, \quad DW \xrightarrow{d} 0,$$

where

$$(22) \quad S_{XY} = \int_0^1 V(t)W(t)dt - \int_0^1 V(t)dt \int_0^1 W(t)dt,$$

$$(23) \quad S_{XX} = \int_0^1 W^2(t)dt - \left(\int_0^1 W(t)dt \right)^2,$$

$$(24) \quad S_{YY} = \int_0^1 V^2(t)dt - \left(\int_0^1 V(t)dt \right)^2,$$

$\sigma_w, \sigma_v > 0$, and the independent Wiener processes $W(t)$ and $V(t)$ arise from the component-wise application of the FCLT. The usual t -statistics does not therefore possess a limit distribution but diverges as the sample size increases; the bias in this test towards the rejection of no relationship based on a given nominal critical value will *increase* with n ; the DW d -statistics converges in probability to zero; and R^2 has a non-degenerated limit distribution as $n \rightarrow \infty$.

A regression model like (19) that appears to find relations that do not really exist is called a spurious regression. The problem here has two elements which co-integration seeks to resolve. Firstly, under the null hypothesis, the model says that y_t is equal to a constant plus a disturbance term, and so the null is false, because y_t is truly generated as a random walk. It is not unusual for tests to reject false null hypotheses even when the alternative is false [20]. Secondly, the standard asymptotic results do not hold when at least one of the regressors is I(1) [76,28]. Our focus is principally on the development of the FCLT for integrated processes, and so we will mention just two other papers here. Sims, Stock and Watson [106] showed in a vector autoregression containing some unit roots and possibly time trends that it is not the case the t -statistics on every parameter involving an I(1) variable follows a non-standard distribution asymptotically: parameters that can be written as coefficients on mean-zero, non-integrated regressors are root n consistent and asymptotically normal. Elliot, Rothenberg and Stock [29] proposed a family of point-optimal tests that dominates the basic unit root tests when a time series has an unknown mean or a linear trend, although Robinson [98] showed if the unit root is nested in a model against a different class of alternatives that include fractionally differenced processes, a standard efficiency theory based on the chi-square distribution is possible (see [107] for a fuller discussion of hypothesis testing.)

We would expect that residuals from a spurious regression such as (19) would be I(1). But if y and x were instead co-integrated variables, we would expect the residuals from

(19) to be $I(0)$. This means that the residuals can be used as the basis of testing for co-integration: we subject the residuals to a unit root test under the null hypothesis that the variables are *not* co-integrated. Limiting distribution theory based on Wiener functionals can be constructed using the FCLT's above for the time series regression with mixtures of integrated processes with an unknown degree of co-integration (see [54],[80],[13],[42]). One problem with the asymptotic theory of non-linear ML estimation in integrated and co-integrated systems is that the classical proofs of consistency based on the uniform convergence of the objective function over the parameter space generally are not applied owing to the objective function diverging at different rates in the parameter space. Even if the consistency can be established, it can be difficult to deduce the limit distribution of the ML estimator from a conventional Taylor series expansion of the score vector. Saikkonen [101] resolved these difficulties by constructing a statistical theory, where, in addition to the consistency, a result on the order of consistency of the estimator of the long-run parameter of the model (pertaining to the co-integrating relation) is available, and the standardized sample information matrix satisfies a suitable stochastic equicontinuity condition. This program of work was completed with two recent papers [102,103] allowing the consistency of the reduced form parameters to be established without assuming the identifiability of the structural parameters and establishing a limit distribution theory without assuming the identifiability of the parameters in the short-run dynamics (i.e., differenced terms and lags of differenced terms not involving the co-integrating relation). These conditions are especially suitable from the point of view of econometric time series.

4. OTHER RECENT WORK

The application of the FCLT to (16) is related to the context of linear regression involving integrated time series. Recently, the focus has switched to non-linear transformations of integrated time series, and an asymptotic theory of inference that is applicable to the non-linear regression has been developed by Park and Phillips [74,75], de Jong [22], de Jong and Wang [23], Pötscher [95], and Berkes and Horváth [6]. Alongside this, related analytic tools on the local time density and hazard functions of the limiting Brownian motion of a standardized integrated process have been developed by Phillips [82,83]. Park and Phillips [74] showed that (16) holds for a class of functions they called "regular" that is wider than the class of continuous functions, although their class does not contain every bounded and measurable function and does not include locally bounded functions such as $T(x) = \log |x|$ or $T(x) = |x|^\alpha (-1 < \alpha < 0)$. de Jong [22] established (16) under local integrability for a different class, covering the two functions above and allowing T to have poles but not encompassing all the regular functions in [74]. The problem was elegantly resolved by Pötscher [95] who established that (16) holds under the condition that the process satisfies the FCLT (as in the "first stage" described above) and then *only* under the additional condition that T is locally integrable. Since the integral in (16) exists almost surely (a.s) and is finite a.s. if and only if T is locally integrable [55, pp. 216–217], the latter condition is minimal. The integral in (16) can be equivalently expressed in terms of the local time:

$$(25) \quad \int_0^1 T(\sigma W(r)) dr = \int_{-\infty}^{\infty} T(\sigma s) L(1, s) ds, \quad a.s.,$$

where $L(t, s)$ is the Brownian local time [96, Ch. 6] which, intuitively, is a spatial density that records the relative sojourn time of the process $W(t)$ at the spatial point s over the time interval $[0, t]$. Equation (25) is known as the occupation times formula [96, p. 224], and it opens up the possibility of using the sojourn times and spatial densities as the basis of a theory of non-parametric co-integrating regression [110].

Suppose (16) holds for a function H , and the function T satisfies $T(\lambda x) = \nu(\lambda)H(x)$ for all $\lambda > 0$ and all $x \in \mathcal{R}$. Then (16) applied to H can be rewritten as

$$(26) \quad (n\nu(n^{-1/2}))^{-1} \sum_{t=1}^n T(x_t) \xrightarrow{d} \int_0^1 H(\sigma W(r)) dr.$$

Expression (26) can still be established if the function T satisfies the equation approximately, in the sense that $\nu(\lambda)^{-1}T(\lambda x) \xrightarrow{L_1} H(x)$ as $\lambda \rightarrow \infty$, thereby providing a basis for convergence results for non-linear functions of the unnormalized integrated processes. This approach is described by Park and Phillips [74] for their “regular” functions and by de Jong and Whang [23] under the conditions in [22], but the result holds more generally under just Pötscher’s basic condition that H is locally integrable. This can then be applied to the context of the non-linear regression with integrated time series [75].

Another topic of recent interest, in the light of the recent observed behaviour in commodity prices and “financial exuberance” in stock markets, is works on mildly explosive processes [87–89,58,36]. This has involved the “moderate deviations from unity” model which is specified as

$$(27) \quad x_t = \rho_n x_{t-1} + \varepsilon_t, \quad (t = 1, \dots, n)$$

$$(28) \quad \rho_n = 1 + \frac{c}{n^\alpha}, \quad \alpha \in (0, 1), \quad c \in \mathcal{R},$$

initialized at some $x_0 = o_p(n^{\alpha/2})$ independent of $\sigma(\varepsilon_1, \dots, \varepsilon_n)$, where ε_t is either a sequence of i.i.d. $(0, \sigma^2)$ random variables with finite ν th moment ($\nu > 2\alpha^{-1}$) [87] or a sequence of weakly dependent random variables [88]. Strictly speaking, since the autoregressive parameter ρ_n is a sequence of the sample size n , (28) is a triangular array $(x_{nt}, 1 \leq t \leq n, n \in \mathcal{N})$, and we adopt the notation x_t and the notation x_0 , for convenience. The idea of model (27) and (28) is “to smooth the passage through unity” relative to (2) such that the roots belong to larger (more moderate) deviations from unity compared with conventional local-to-unity roots [77,11]. The boundary value as $\alpha \rightarrow 1$ includes the conventional local-to-unity case, and the boundary value as $\alpha \rightarrow 0$ includes the stationary or explosive AR(1) process depending on the value of c . Phillips and Magdalinos [87,88] combined a functional law to a diffusion on $D[0, \infty)$ with a central limit law to a Gaussian random variable to establish the limit distribution of the centered and normalized serial correlation coefficient

$$(29) \quad \hat{\rho}_n - \rho = \left(\sum_{t=1}^n x_{t-1} \right)^2 \sum_{t=1}^n x_{t-1} \varepsilon_t.$$

For $c < 0$, they established an $n^{(1+\alpha)/2}$ rate of convergence to the asymptotic normality, bridging the \sqrt{n} and n convergence rates for the stationary and conventional local-to-unity cases. For $c > 0$, the limit distribution is a Cauchy one and is invariant to both the distribution and the dependence structure of innovations. The rate of convergence is $n^\alpha \rho_n^n$, bridging the n and $(1+c)^n$ convergence rates for the conventional local-to-unity and explosive cases. Paper [88] contains possibly the first general invariance principle in statistics that is applicable to explosive processes.

This current program of work seems to offer the immediate challenge of either designing a model offering a *uniform* limit theory for autoregression, or at least constructing an optimality theory that is applied uniformly across parameter values. Shiryaev and Spokoiny [104] have already shown that the two distributions in (12) can be reduced to one—a centered Gaussian variable with variance one—by considering a *sequential* maximum likelihood estimator. Phillips and Han [85], in an approach based on taking the

first differences, offer the estimates that have virtually no finite sample bias and are not sensitive to initial conditions, and the Gaussian central limit theory is applied as the autoregressive coefficient passes through unity with uniform root n convergence (see also [36]). There are still, of course, difficulties that need to be overcome [51,92–94]. It should also be borne in mind that, while we have focused on the econometric literature that has been spawned by the application of the FCLTs based on the Skorokhod topology, this involves an asymptotic approach, and the true objective remains characterizing the finite sample properties of an appropriate statistics.

BIBLIOGRAPHY

1. Abadir K.M., *A distribution generating equation for unit-root statistics*, Oxford Bulletin of Economics and Statistics **54** (1992), 305–323.
2. Abadir K.M., *On the asymptotic power of unit root tests*, Economet. Theory **9** (1993), 189–221.
3. Anderson T.W., *On the asymptotic distributions of estimates of stochastic difference equations*, Annals of Mathematical Statistics **30** (1959), 676–687.
4. Ango Nze P., Doukhan P., *Weak dependence: models and applications to econometrics*, Economet. Theory **20** (2004), 995–1045.
5. Beare B.K., *A new mixing condition*, University of Oxford, Department of Economics, Discussion Paper No. 348 (2007).
6. Berkes I., Horváth L., *Convergence of integral functionals of stochastic processes*, Economet. Theory **22** (2006), 304–322.
7. Beveridge S. Nelson C.R., *A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to measurement of the “business cycle”*, J. Monetary Economics **7** (1981), 151–174.
8. Billingsley P., *Convergence of Probability Measures*, Wiley, 1968.
9. Box G.E.P., Jenkins G.M., *Time Series Analysis: Forecasting and Control*, Holden–Day, 1970.
10. Blough S.R., *The relationship between power and level for generic unit root tests in finite samples*, J. Applied Economet. **7** (1992), 295–308.
11. Chan N.H., Wei C.Z., *Limiting distributions of least squares estimates of unstable autoregressive processes*, Annals of Statistics **16** (1988), 367–401.
12. Chan N.H., Tran L.T., *On the first order autoregressive process with infinite variance*, Economet. Theory **5** (1989), 354–362.
13. Chang Y., Phillips P.C.B., *Time series regression with mixtures of integrated processes*, Economet. Theory **11**, 1033–1094.
14. Cooper R.W., *Econometric models of cyclical behaviour*, National Bureau of Economic Research Studies in Income and Wealth No. 36, Columbia University Press, 1972.
15. Davidson J., *Stochastic Limit Theory*, vol. 1994, Oxford University Press.
16. Davidson J., *Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes*, Journal Economet. **106** (2002), 243–269. (Corrigendum: J. Economet. vol 110 (2002) 103–104.
17. Davidson J., *Asymptotic methods and functional central limit theory*, Palgrave Handbook of Econometrics: Vol. 1, Econometric Theory, Chapter 5, ed. T.C. Mills and K. Patterson, Palgrave Macmillan, 2006.
18. Davidson J., *When is a time series $I(0)$?* (2007, mimeo, University of Exeter).
19. Davidson J., de Jong R.M., *The functional central limit theorem and weak convergence to stochastic integrals II: fractionally integrated processes*, Economet. Theory **16** (2000), 643–666.
20. Davidson R., MacKinnon J.G., *Implicit alternatives and the local power of test statistics*, Econometrica **60** (1987), 1035–1329.
21. de Jong R.M., *Addendum to: “Asymptotics for nonlinear transformations of integrated time series”*, Economet. Theory **20** 627–635.
22. de Jong R.M., Davidson J., *The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent processes*, Economet. Theory **16** (2000), 621–642.
23. de Jong R.M., Wang C.-H., *Further results on the asymptotics for nonlinear transformations of integrated time series*, Economet. Theory **21** (2005), 413–430.
24. Dickey D.A., Fuller W.A., *Distribution of the estimators for autoregressive time series with a unit root*, J. Am. Statist. Assoc. **74** 427–431.
25. Dickey D.A., Fuller W.A., *Likelihood ratio statistic for autoregressive time series with a unit root*, Econometrica **49** (1981), 1057–1072.

26. Donsker M.D., *An invariance principle for certain probability limit theorems*, Memoirs of the American Mathematical Society **6** (1951), 1–12.
27. Doukhan P., Louhichi S., *A new weak dependence condition and applications to moment inequalities*, Stochastic Process Appl. **84** (1999), 313–342.
28. Durlauf S.N., Phillips P.C.B., *Trends versus random walks in time series analysis*, Econometrica **56** (1988), 1333–1354.
29. Elliot G., Rothenberg T., Stock J.H., *Efficient tests for an autoregressive unit root*, Econometrica **64** **813–836**.
30. Engle R.F., Granger C.W.J., *Co-integration and error correction: representation estimation and testing*, Econometrica **55** (1987), 251–276.
31. Erdős P., Kac M., *On certain limit theorems in the theory of probability*, Bulletin of the American Mathematical Society **52** (1946), 292–302.
32. Faust J., *Near observational equivalence and theoretical size problems with unit root tests*, Economet. Theory **12** (1996), 724–731.
33. Faust J., *Conventional confidence intervals for points on spectrum have confidence level zero*, Econometrica **67** (1999), 629–637.
34. Gil-Alaña L., Robinson P.M., *Testing of unit root and other nonstationary hypotheses in macroeconomic time series*, J. Economet. **80** (1997), 241–268.
35. Grenander U., Rosenblatt M., *Statistical Analysis of Stationary Time Series*, Wiley, 1958.
36. Giraitis L., Phillips P.C.B., *Uniform limit theory for stationary autoregression*, J. Time Ser. Anal. **27** (2006), 51–60. (Corrigendum: J. Time Ser. Anal. vol 27 no. 6 (2006) i-ii.
37. Granger C.W.J., *Some properties of time series data and their use in econometric model specification*, J. Economet. **16** (1981), 121–130.
38. Granger C.W.J., *Co-integrated variables and error-correcting models* (1983), unpublished, University of California, San Diego, Discussion Paper 83–13.
39. Granger C.W.J., *Developments in the study of cointegrated economic variables*, Oxford Bulletin of Economics and Statistics **48** (1986), 213–228.
40. Granger C.W.J., Newbold P., *Spurious regressions in econometrics*, J. Economet. **2** (1974), 111–120.
41. Hall R.E., *Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence*, J. Political Economy **96** (1978), 971–987.
42. Hamilton J.D., *Time Series Analysis*, Princeton, 1994.
43. Hannan E.J., *Multiple Time Series*, Wiley, 1970.
44. Hendry D.F., *Econometric modelling with cointegrated variables: an overview*, Oxford Bulletin of Economics and Statistics **48** (1986), 201–212.
45. Hendry D.F., Juselius K., *Explaining cointegration models: part I*, Energy Journal **21** (2000), 1–42.
46. Hendry D.F., Juselius K., *Explaining cointegration models: part II*, Energy Journal **22** (2001), 75–120.
47. Hendry, D.F. Pagan A.R., Sargan J.D., Griliches Z., Intriligator M.D. (eds.) *Handbook of Econometrics: Vol. II*, Elsevier.
48. Herndorf N., *The invariance principle for φ -mixing sequences*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **63** (1983), 97–108.
49. Herndorf N., *A functional central limit theorem for ρ -mixing sequences*, J. Multivariate Anal. **15** (1984), 141–146.
50. Herndorf N., *A functional central limit theorem for weakly dependent sequences of random variables*, Ann. Probab. **12** (1984), 141–153.
51. Jeganathan P., *On asymptotic inference in linear cointegrated time series systems*, Economet. Theory **13** (1995), 692–745.
52. Jeganathan P., *Convergence of functionals of sums of r.v.'s to local times of fractional stable motions*, Ann. Probab. **32** (2004), 1771–1795.
53. Johansen S. *Statistical analysis of cointegration vectors*, J. Econ. Dynam. Contr. **12** (1988), 231–254.
54. Johansen S., *Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models*, Econometrica **59** (1991), 1551–1580.
55. Karatzas I., Shreve S.E., *Brownian Motion and Stochastic Calculus*, Springer, 1991.
56. Koul H., Pflug G., *Weakly adaptive estimators in explosive autoregression*, Ann. Statist. **18** (1990), 939–960.
57. Leeb H., Pötscher B.M., *The variance of a integrated process need not diverge to infinity and related results on the partial sums of stationary processes*, Economet. Theory **17** (2001), 671–685.

58. Magdalinos T., *On the inconsistency of the unrestricted estimator of the information matrix near a unit root*, *Economet. J.* **10** (2007), 245–262.
59. Mann H.B., Wald A., *On the statistical treatment of linear stochastic difference equations*, *Econometrica* **11** (1943), 173–200.
60. Marinucci D., Robinson P.M., *Alternative forms of fractional Brownian motion*, *J. Statist. Plann. Inference* **80** (1999), 111–122.
61. Marinucci D., Robinson P.M., *Weak convergence of multivariate fractional processes*, *Stochastic Processes and their Applications* **86** (2000), 103–120.
62. Marinucci D., Robinson P.M., *Narrow-band analysis of nonstationary processes*, *Annals of Statistics* **29** (2001), 947–986.
63. Marsh P., *The available information for invariant tests of a unit root*, *Economet. Theory* **23** (2007), 686–710.
64. McCrorie J.R., *Representations of the moments of the Dickey–Fuller distribution*, preprint.
65. McLeish D.L., *Invariance principles for dependent variables*, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** (1975), 165–178.
66. McLeish D.L., *On the invariance principle for nonstationary mixingales*, *Ann. Probab.* **5** (1977), 616–621.
67. Mijnhoeer J., *Asymptotic inference for AR(1) processes with (nonnormal) stable errors*, *J. Mathematical Sciences* **83** (1997), 401–406.
68. Mikulski P.W., Monsour M.J., *Optimality of the maximum likelihood estimator in first-order autoregressive processes*, *J. Time Ser. Anal.* **12** (1991), 237–253.
69. Müller U.K., *The impossibility of consistent discrimination between $I(0)$ and $I(1)$ processes*, *Economet. Theory* **24** (2008 forthcoming).
70. Naylor T.H., Seaks T.G., Wichern D.W., *Box–Jenkins models: an alternative to econometric models*, *International Statistical Review* **40** (1972), 123–137.
71. Nelson C.R., Plosser C.I., *Trends and random walks in macroeconomic time series*, *Journal of Monetary Economics* **10** (1982), 139–162.
72. Paparoditis E., Politis D.N., *Bootstrapping unit root tests for autoregressive time series*, *J. Am. Statist. Assoc.* **100** (2005), 545–553.
73. Park J.Y., *Bootstrap unit root tests*, *Econometrica* **71** (2003), 1845–1895.
74. Park J.Y., Phillips P.C.B., *Asymptotics for nonlinear transformations of integrated time series*, *Economet. Theory* **15** (1999), 269–298.
75. Park J.Y., Phillips P.C.B., *Nonlinear regressions with integrated time series*, *Econometrica* **69** (2001), 117–161.
76. Phillips P.C.B., *Understanding spurious regressions in econometrics*, *J. Economet.* **33** (1986), 322–340.
77. Phillips P.C.B., *Time series regression with a unit root*, *Econometrica* **55** (1987), 277–301.
78. Phillips P.C.B., *Time series regression with a unit root and infinite variance errors*, *Economet. Theory* **6** (1990), 44–62.
79. Phillips P.C.B., *Optimal inference in cointegrated systems*, *Econometrica* **59** (1991), 283–306.
80. Phillips P.C.B., *Fully modified least squares and vector autoregression*, *Econometrica* **63** (1995), 1023–1078.
81. Phillips P.C.B., *New tools for understanding spurious regressions*, *Econometrica* **66** (1998), 1299–1325.
82. Phillips P.C.B., *Descriptive econometrics for non-stationary time series with empirical illustrations*, *Journal of Applied Econometrics* **16** (2001), 389–413.
83. Phillips P.C.B., *Econometric analysis of Fisher’s equation*, *Am. J. Econ. Sociology* **64** (2005), 125–168.
84. Phillips P.C.B., Durlauf S.N., *Multiple time series regression with integrated processes*, *Review of Economic Studies* **53** (1986), 473–495.
85. Phillips P.C.B., Han, *Gaussian inference in AR(1) time series with or without a unit root*, *Economet. Theory* **24** (2008 forthcoming).
86. Phillips P.C.B., Loretan M., *Estimating long-run economic equilibria*, *Review of Economic Studies* **57** (1991), 99–125.
87. Phillips P.C.B., Magdalinos T., *Limit theory for moderate deviations from a unit root*, *J. Economet.* **136** (2007), 115–130.
88. Phillips P.C.B., Magdalinos T., Phillips G.D.A., Tzavalis E. (eds.) *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*, Cambridge University Press, 2007.
89. Phillips P.C.B., Magdalinos T., *Limit theory for explosively cointegrated systems*, *Economet. Theory* **24** (2008, forthcoming).

90. Phillips P.C.B., Perron P., *Testing for a unit root in time series regressions*, *Biometrika* **75** (1988), 335–346.
91. Phillips P.C.B., Solo V., *Asymptotics for linear processes*, *Ann. Statist.* **20** (1992), 971–1001.
92. Ploberger W., *A complete class of tests when the likelihood is locally asymptotically quadratic*, *J. Economet.* **118** (2004), 67–94.
93. Ploberger W., *Admissible and nonadmissible tests in unit-root-like situations*, *Economet. Theory* **24** (2008 forthcoming).
94. Pötscher B.M., *Lower risk bounds and properties of confidence sets for ill-posed estimation problems with applications to spectral density and persistence estimation unit roots and estimation of long memory parameters*, *Econometrica* **70** (2002), 1035–1065.
95. Pötscher B.M., *Nonlinear functions and convergence to Brownian motion: beyond the continuous mapping theorem*, *Economet. Theory* **20** (2004), 1–22.
96. Revuz D., Yor M., *Continuous Martingales and Brownian Motion* (1999, Springer).
97. Robinson P.M., *Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regressions*, *J. Economet.* **47** (1991), 57–84.
98. Robinson P.M., *Efficient tests of nonstationary hypotheses*, *J. Am. Statist. Assoc.* **89** (1994), 1420–1437.
99. Robinson P.M., Hualde J., *Cointegration in fractional systems with unknown integration orders*, *Econometrica* **71** (2003), 1727–1766.
100. Robinson P.M., Yajima Y., *Determination of cointegrating rank in fractional systems*, *J. Economet.* **106** (2002), 217–241.
101. Saikkonen P., *Problems with the asymptotic theory of maximum likelihood estimation in integrated cointegrated systems*, *Economet. Theory* **11** (1995), 888–911.
102. Saikkonen P., *Consistent estimation in cointegrated vector autoregressive models with nonlinear time trends in cointegrating relations*, *Economet. Theory* **17** (2001), 296–326.
103. Saikkonen P., *Statistical inference in cointegrated vector autoregressive models with nonlinear time trends in cointegrating relations*, *Economet. Theory* **17** (2001), 327–356.
104. Shiryaev A.N., Spokoiny V.G., *On sequential estimation of an autoregressive parameter*, *Stochastics and Stochastic Reports* **60** (1997), 219–240.
105. Skorokhod A.V., *Limit theorems for stochastic processes*, *Theory of Probability and its Applications* **1** (1956), 261–290.
106. Sims C.A., J.H. Stock M.W., *Watson inference in linear time series models with some unit roots*, *Econometrica* **58** (1990), 113–144.
107. Stock J.H., *Handbook of Econometrics*, vol. 4, Elsevier, 1994.
108. Stock J.H., Watson M.W., *Testing for common trends*, *J. Am. Statist. Assoc.* **83** (1988), 1097–1107.
109. Tanaka K., *The nonstationary fractional unit root*, *Econometric Theory* **15** (1999), 549–582.
110. Wang Q., Phillips P.C.B., *Asymptotic theory for local time density estimation and nonparametric cointegrating regression* (2006), Cowles Foundation, Discussion Paper No. 1594.
111. Watson M.W., *R.F. Engle, D.L. McFadden (eds.) Handbook of Econometrics, Vol. 4*, Elsevier, 1994.
112. White J.S., *The limiting distribution of the serial correlation coefficient in the explosive case*, *Annals of Mathematical Statistics* **29** (1958), 1188–1197.
113. Wooldridge J.M., White H., *Some invariance principles and central limit theorems for dependent heterogeneous processes*, *Economet. Theory* **4** (1988), 210–230.

SCHOOL OF ECONOMICS AND FINANCE, UNIVERSITY OF ST. ANDREWS, CASTLECLIFFE, THE
 SCORES, ST. ANDREWS, KY16 9AL, UNITED KINGDOM
E-mail: mccrorie@st-andrews.ac.uk