Asymptotic Properties of Growth Rates

Christine Choirat* Raffaello Seri†

February 18, 2005

Abstract

Consider a series $(X_t)_t$ such that the ratio of two successive values $\frac{X_{t+1}}{X_t}$ is a stationary variable. We introduce two growth rates, the *average growth rate* $\bar{R}$ and the *long-term growth rate* ρ. The behavior of the average growth rate is standard, since it is always asymptotically normal, with a variance that can be computed. We show that the *average growth rate bias* observed by Day and Huang (2001) and Huang (2004) in dynamical systems holds in this case too, since when the process $(X_t)_t$ is stationary $\bar{R}$ is strictly positive. On the other hand, the asymptotic behavior of the long-term growth rate ρ deserves more attention. Its estimator is usually asymptotically normal, but it is non-standard when $(X_t)_t$ is stationary. We obtain the rate of convergence of the finite-sample distribution to the asymptotic one in the form of a Berry-Essen bound that is shown to depend on the memory of the process. Moreover, we provide an algorithm based on kernel non-parametric density estimation to evaluate the asymptotic distribution of the growth rate, whose rate of convergence is bounded from above.

Acknowledgements. We are glad to thank N. Bonneuil, A. Fischlin, P. Haccou, J. Hofbauer, C. Jacob, S. Johansen and J.A.J. Metz for helpful comments.

1 Introduction

Consider a series $(X_t)_t$ such that the ratio of two successive values $\frac{X_{t+1}}{X_t}$ is almost a stationary variable (indeed, in part of our limit results, we will allow for this ratio to asymptotically mean stationary). It is customary in the econometric study of nonstationary random processes to consider processes such that $Y_{t+1} - Y_t$ is a stationary random variable (sometimes even a white noise plus a constant). We prefer another formulation in which it is the ratio of two successive values, i.e. $\frac{X_{t+1}}{X_t}$, that has this property. This entails no loss of generality

*Dipartimento di Economia, Universitá degli Studi dell'Insubria, Via Ravasi 2, 21100 Varese, Italy, email: `cchoirat@eco.uninsubria.it`

†Dipartimento di Economia, Universitá degli Studi dell'Insubria, Via Ravasi 2, 21100 Varese, Italy, email: `rseri@eco.uninsubria.it`

since a logarithmic transformation on $(X_t)_t$ leads to a $(Y_t)_t$ that is stationary in difference, and an exponential transformation on $(Y_t)_t$ leads to a $(X_t)_t$ that is stationary in ratio. This preference is motivated by the fact that the present theory has applications that go well beyond the economic domain, since it invests ecology, biology, chemistry, physics and so on, i.e. any domain in which it is of interest to test whether a phenomenon is growing or not.

We introduce two growth rates defined in term of the original process $(X_t)_t$ as follows. The *average growth rate* is given by:

$$\bar{R} \triangleq \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T-1} \frac{X_{t+1} - X_t}{X_t},$$

and the *long-term growth rate* is given by:

$$\rho \triangleq \lim_{T\to\infty} \left(\frac{X_T}{X_1} \right)^{\frac{1}{T}} - 1.$$

Remark that also ρ can be written as a sum:

$$\ln(1+\rho) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T-1} \ln\left(\frac{X_{t+1}}{X_t} \right)$$

Even if it is usually supposed that these two quantities have a similar behavior (i.e. that $\frac{X_{t+1}-X_t}{X_t} \approx \ln \frac{X_{t+1}}{X_t}$, see e.g. Hamilton, 1994, p. 438), we show that this is not always the case. As explained above, we consider the situation in which the process $(X_t)_t$ has a stationary instantaneous growth rate (this means that the ratio of successive values of the process form a stationary sequence). Indeed, $\bar{R}$ is a well-defined limit (that is a constant in the ergodic case and more generally a random variable in the stationary case) that can be positive or negative. Strangely enough, when the process $(X_t)_t$ is stationary, $\bar{R}$ is strictly positive: according to the usual interpretation of a growth rate this gives the impression that the process is increasing! This phenomenon has been apparently first observed by Day and Huang (2001) and Huang (2004) in dynamical systems, where it has been called *average growth rate bias*. Huang (2004, p. 47) states that the phenomenon should hold also for stochastic processes: we show that this is indeed the case. Notwithstanding this strange phenomenon, the behavior of $\bar{R}$ is standard, since it is asymptotically normal both under stationarity of $(X_t)_t$ and of $\left(\frac{X_{t+1}}{X_t}\right)_t$, with a variance that can be computed using some econometric techniques. The fact that $\bar{R}$ is positive for stationary processes leads to the impossibility of basing a test of stationarity based on this definition of the growth rate.

On the other hand, the asymptotic behavior of ρ deserves more attention. When $\left(\frac{X_{t+1}}{X_t}\right)_t$ is stationary (but $(X_t)_t$ is not), ρ is different from zero, while if $(X_t)_t$ is stationary $\rho = 0$. This points at ρ as a good candidate for an indicator of growth. The main problem is the asymptotic distribution of the estimator of

ρ: when $\left(\frac{X_{t+1}}{X_t}\right)_t$ is stationary (but $(X_t)_t$ is not), the asymptotic distribution is normal, while it is non-standard (or normal with degenerate null variance) when $(X_t)_t$ is stationary. This implies also that a confidence region for ρ when the true value is $\rho = 0$ appears to be misleadingly precise and often not centered around zero (indeed, the variance seems to converge to 0 faster than the estimator). On the other hand, when $\rho = 0$, the distribution is non-standard and depends dramatically on the marginal distribution of the process $(X_t)_t$. We obtain the rate of convergence of the finite-sample distribution to the asymptotic one in the form of a Berry-Essen bound that is shown to depend on the integrability properties of $\ln X_t$, on the density of $\ln X_t$ and, what is more relevant, on the memory of the process (when X_t is multidimensional, also the dimension of the space is important). The result uses some coupling Theorems for dependent processes. Moreover, we provide an algorithm based on kernel non-parametric density estimation to evaluate the asymptotic distribution of the growth rate, whose rate of convergence is bounded from above. When the process $(X_t)_t$ is stationary the procedure seems to work well. On the other hand, when $\left(\frac{X_{t+1}}{X_t}\right)_t$ is stationary (but $(X_t)_t$ is not, i.e. under the alternative), the density of the limiting distribution as evaluated by the kernel estimator becomes flatter and flatter as T increases and the estimator of ρ moves towards the queue of the distribution, leading to a reject of the null hypothesis. This phenomenon is highly desirable and reasonable but still lacks a mathematical explanation.

2 The Results

In this Section, we introduce two different definitions of growth rates that have been presented in the literature, the average growth rate and the long-term growth rate. Strangely enough, these coefficients are defined by a limit and not by a closed form expression. Therefore, in the following Sections, we derive these expressions under an assumption of asymptotic mean stationarity, and we show how they change when some additional hypotheses are imposed. Then we derive the asymptotic distributions of these coefficients, and we give methods to calculate them. These results can be implemented to derive confidence intervals on the value of the growth rates.

We define the *instantaneous growth rate* R_t^i by the equality:

$$R_t^i \triangleq \begin{cases} 0 & \text{when } X_t^i = 0 \\ \frac{X_{t+1}^i - X_t^i}{X_t^i} & \text{else} \end{cases}$$

Therefore we have:

$$X_{t+1}^i = \left(1_{\{X_t^i > 0\}} + R_t^i\right) \cdot X_t^i;$$

where the product has to be read as 0 when $X_t^i = 0$ in order to avoid indeterminacy. To avoid this situation, we will suppose in the following that X_t^i is always strictly positive.

From the previously defined instantaneous growth rate R_t^i, three different indicators of growth can be derived. We recall the definition of the *average growth rate*:

$$\bar{R}^i \triangleq \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} R_t^i = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \left(\frac{X_{t+1}^i - X_t^i}{X_t^i} \cdot 1_{\{X_t^i>0\}} \right) \tag{2.1}$$

and the *long-term growth rate*:

$$\rho^i \triangleq \lim_{T\to\infty} \left[\prod_{t=0}^{T-1} \left(1_{\{X_t^i>0\}} + R_t^i \right) \right]^{1/T} -1 = \lim_{T\to\infty} \left[\prod_{t=0}^{T-1} \left(\frac{X_{t+1}^i}{X_t^i} \cdot 1_{\{X_t^i>0\}} \right) \right]^{1/T} -1. \tag{2.2}$$

Remark that $\rho^i = \exp\left(r^i\right) - 1$ where

$$r^i \triangleq \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \ln\left(1_{\{X_t^i>0\}} + R_t^i \right) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \ln\left(\frac{X_{t+1}^i}{X_t^i} \cdot 1_{\{X_t^i>0\}} \right)$$

is the *logarithmic growth rate*. Moreover, the two coefficients ρ^i and $\bar{R}^i$ are linked by the inequality:

$$1 + \rho^i \leq \exp\left(\bar{R}^i\right),$$

with equality iff X_t^i is equal to a strictly positive constant for almost every t.

2.1 The Average Growth Rate $\bar{R}^i$

In this Section, we investigate the asymptotic properties of the average growth rate. The first result states a condition under which the average growth rate has a well-defined limit and derives a formula for this limit.

Theorem 2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbf{X}_t)_{t\in\mathbb{N}}$ a strictly positive process. Suppose, moreover, that $\left(R_t^i\right)_{t=0,1,\ldots}$ is asymptotically mean stationary with asymptotic mean $\mathbb{P}^*$. Then, $\mathbb{P}-$ as and $\mathbb{P}^*-$ as:*

$$\bar{R}^i(\omega) = \mathbb{E}^*\left[\frac{X_0^i(S\cdot)}{X_0^i(\cdot)} |\mathcal{I} \right](\omega) - 1.$$

If the asymptotic mean $\mathbb{P}^*$ is ergodic, then, $\mathbb{P}-$ as and $\mathbb{P}^*-$ as:

$$\bar{R}^i = \mathbb{E}^*\left[\frac{X_0^i(S\cdot)}{X_0^i(\cdot)} \right] - 1.$$

Remark 2.1. (i) *The limit of the average growth rate in the nonergodic case is generally a random variable.*
(ii) *The result is stated in a probabilistic framework. However, Birkhoff Ergodic Theorem also holds for deterministic dynamical systems (see Katok and Hasselblatt, 1995, p. 136, and Lasota and Mackey, 1994, p. 63). In that*

case, though, the limit is seldom interpreted as a conditional expectation operator since conditional probabilities are not a common tool of dynamical systems theorists. In these cases, the conditional expectation is substituted by an unspecified $\mathcal{I}-$measurable function. Moreover, Theorem B.2 that we state in Appendix B (taken from Choirat et al.*, 2003, 2005) does not require integrability of the random variables.*

In what follows, we use Doukhan, Massart and Rio's (1994) Central Limit Theorem for strongly mixing processes (see Theorem B.4) in order to obtain the asymptotic distribution of the average growth rate.

Theorem 2.2. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbf{X}_t)_{t\in\mathbb{N}}$ a strictly positive process. Suppose that $\left(R_t^i\right)_{t=0,1,\ldots}$ is an $\mathcal{A}-$measurable stationary strongly mixing transformation and suppose the following hypotheses hold:*
***(i)** $X_0^i(\cdot)$ is $\mathbb{P}-\mathsf{as}$ strictly positive;*
***(ii)** $\mathbb{E}\left(\frac{X_{t+1}^i}{X_t^i}\right) = \mu$, $\mathbb{E}\left(\frac{X_{t+1}^i}{X_t^i} - \mu\right)^2 = \sigma^2$, $\mathbb{E}\left[\left(\frac{X_{t+1}^i}{X_t^i} - \mu\right)\left(\frac{X_{t+s+1}^i}{X_{t+s}^i} - \mu\right)\right] = \sigma^2 \cdot \gamma_s$;*
***(iii)** for some $r > 2$, $\mathbb{E}\left|\frac{X_{t+1}^i}{X_t^i} - \mu\right|^r < \infty$;*
***(iv)** for the same $r > 2$, $\sum_{t=1}^{+\infty} t^{2/(r-2)}\alpha(t) < \infty$;*
then $\sum_{t=1}^{+\infty} |\gamma_t| < \infty$ and:

$$\sqrt{T}\left[\frac{1}{T}\sum_{t=0}^{T-1} R_t^i - \bar{R}^i\right] \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2 \cdot \left(1 + 2 \cdot \sum_{s=1}^{+\infty}\gamma_s\right)\right).$$

Remark 2.2. (i) *The previous Theorem can be easily generalized to the multivariate case using the Cramér-Wold device: for random vectors $\mathbf{Z}$ and $\mathbf{Z}_1, \mathbf{Z}_2, \ldots$ in $\mathbb{R}^d$, $\mathbf{Z}_n \xrightarrow{\mathcal{D}} \mathbf{Z}$ if and only if $\lambda^\mathsf{T}\mathbf{Z}_n \xrightarrow{\mathcal{D}} \lambda^\mathsf{T}\mathbf{Z}$ for all $\lambda \in \mathbb{R}^d$, where $\mathbf{A}^\mathsf{T}$ denotes the transpose of the matrix $\mathbf{A}$. Remark, however, that in this case the conditions should be applied to the whole vector process $(\mathbf{X}_t)_{t\in\mathbb{N}}$.*
(ii) *The Central Limit Theorem on which we base our result (see Doukhan, Massart and Rio, 1994) is not the most general result in the literature, but it is a good compromise between generality and simplicity. Indeed, Rio's (1995) Central Limit Theorem holds for nonnecessarily stationary strongly mixing processes and provides also a rate of convergence to normality: however, the verification of its conditions appear to be quite involved.*
(iii) *For specific models, the conditions in the statement of the Theorem can be checked on a case-by-case basis.*
(iv) *This Theorem can be generalized to deterministic dynamical systems using, for example, the* CLT*'s in Liverani (1996).*

Now, we can consider testing based on the average growth rate. The previous results can be used in order to build a confidence interval for $\bar{R}^i$, once we have an estimator of the asymptotic variance $\sigma^2 \cdot \left(1 + 2 \cdot \sum_{n=1}^{+\infty}\rho_n\right)$. The estimation of this variance is not a simple problem but we can use some results, such as the ones in Andrews (1991), Andrews and Monahan (1992), de Jong and Davidson (2000) and Newey and West (1987, 1994). A review of the methods is in Davidson (2000, Section 9.4.3).

2.2 The Long-term Growth Rate ρ^i

In this Section, we investigate the asymptotic properties of the long-term growth rate.

2.2.1 Limit Theorems

In the following we will write this coefficient as:

$$\rho_T^i(\omega) \triangleq \left[\prod_{t=0}^{T-1}\left(1+R_t^i\right)\right]^{1/T} - 1 = \left[\frac{X_T^i}{X_0^i}\right]^{1/T} - 1.$$

It should be noticed that Remark 2.1 still holds.

Theorem 2.3. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbf{X}_t)_{t\in\mathbb{N}}$ a strictly positive process. Suppose, moreover, that $\left(\ln\left(\frac{X_{t+1}^i}{X_t^i}\right)\right)_{t=0,1,\ldots}$ is asymptotically mean stationary with asymptotic mean $\mathbb{P}^*$. Then, $\mathbb{P}-$as and $\mathbb{P}^*-$as:*

$$\rho^i(\omega) = \exp\left\{\mathbb{E}^*\left[\ln\left(\frac{X_{t+1}^i}{X_t^i}\right)|\mathcal{I}\right]\right\} - 1. \tag{2.3}$$

2.2.2 Asymptotic Distribution of ρ_T^i

In the following we derive the asymptotic distribution of ρ_T^i under some conditions. It will be apparent that ρ_T^i has two very different behaviors if $\left(X_t^i\right)_{t=0,1,\ldots}$ is stationary (Theorem 2.7) or if it is not (Theorem 2.4).

Theorem 2.4. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbf{X}_t)_{t\in\mathbb{N}}$ a strictly positive process. Suppose, moreover, that $\left(R_t^i\right)_{t=0,1,\ldots}$ is an $\mathcal{A}-$measurable stationary strongly mixing transformation and suppose the following hypotheses hold:*
(i) $\mathbb{E}\ln\left(1+R_t^i\right)=\mu$, $\mathbb{E}\left(\ln\left(1+R_t^i\right)-\mu\right)^2=\sigma^2$, $\mathbb{E}\left[\left(\ln\left(1+R_t^i\right)-\mu\right)\left(\ln\left(1+R_{t+s}^i\right)-\mu\right)\right]=\sigma^2\cdot\rho_s$;
(ii) *for some* $r>2$, $\mathbb{E}\left|\ln\left(1+R_t^i\right)-\mu\right|^r<\infty$;
(iii) *for the same* $r>2$, $\sum_{t=1}^{+\infty} t^{2/(r-2)}\alpha(t)<\infty$;
(iv) $\left(X_t^i\right)_{t=0,1,\ldots}$ *is not a stationary process;*
then $\sum_{t=1}^{\infty}|\rho_t|<\infty$ *and:*

$$\sqrt{T}\cdot\ln\left(\frac{1+\rho_T^i}{1+\rho^i}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\cdot\left(1+2\cdot\sum_{s=1}^{+\infty}\rho_s\right)\right).$$

Remark 2.3. *From a formal point of view, the Theorem holds also without Hypothesis* **(v)** *but it loses part of its interest. Indeed, when the process* $\left(X_t^i\right)_{t=0,1,\dots}$ *is stationary,* $\sqrt{T}\cdot\ln\left(\frac{1+\rho_T^i}{1+\rho^i}\right)=\sqrt{T}\cdot\ln\left(1+\rho_T^i\right)$ *converges to a degenerate Gaussian random variable with* $\sigma^2\cdot\left(1+2\cdot\sum_{s=1}^{+\infty}\rho_s\right)=0$*. Theorem 2.7 shows that in this case the asymptotic distribution of* ρ_T^i *is nonstandard.*

As corollaries of Theorems 2.1 and 2.3, we get the following results. We recall that, given any sequence of events $A_1, A_2, \ldots \in \mathcal{A}$, (A_n) is said to hold *ultimately* if it holds for any but a finite number of indexes.

Corollary 2.5. *Let* $(\Omega,\mathcal{A},\mathbb{P})$ *be a probability space and* $(\mathbf{X}_t)_{t\in\mathbb{N}}$ *a strictly positive process. Suppose that* $\left(X_t^i\right)_{t=0,1,\dots}$ *is asymptotically mean stationary with asymptotic mean* $\mathbb{P}^*$*. Then we have* $\bar{R}^i(\omega)=0$ *if the process is ultimately constant, while* $\bar{R}^i(\omega)$ *is strictly positive otherwise. Moreover, if* $\mathbb{P}^*$ *is ergodic we have* $\bar{R}^i=0$ *if the proces is* $\mathbb{P}-\mathsf{as}$ *ultimately constant, while* $\bar{R}^i$ *is strictly positive otherwise.*

Corollary 2.6. *Let* $(\Omega,\mathcal{A},\mathbb{P})$ *be a probability space and* $(\mathbf{X}_t)_{t\in\mathbb{N}}$ *a strictly positive process. Suppose, moreover, that* $\left(X_t^i\right)_{t=0,1,\dots}$ *is asymptotically mean stationary with asymptotic mean* $\mathbb{P}^*$*. Then we have* $\rho_i=0$*.*

The following Theorem gives the asymptotic distribution of ρ_T^i.

Theorem 2.7. *Let* $(\Omega,\mathcal{A},\mathbb{P})$ *be a probability space and* $(\mathbf{X}_t)_{t\in\mathbb{N}}$ *a strictly positive process. Suppose, moreover, that* $\left(X_t^i\right)_{t=0,1,\dots}$ *is a* $2-\alpha-$*mixing stationary process (see Appendix B for definitions) and that* X_0^i *is* $\mathbb{P}-\mathsf{as}$ *strictly positive. Then:*

$$T\cdot\ln\left(1+\rho_T^i\right)\xrightarrow{\mathcal{D}}Z'-Z'', \tag{2.4}$$

where $Z'\overset{\mathcal{D}}{=}Z''\overset{\mathcal{D}}{=}\ln X_0^i$ *and* Z' *is independent of* Z''*.*

Remark 2.4. (i) *The same result holds in the multivariate case:*

$$T\cdot\begin{bmatrix}\ln\left(1+\rho_T^1\right)\\ \vdots\\ \ln\left(1+\rho_T^k\right)\end{bmatrix}\xrightarrow{\mathcal{D}}\mathbf{Z}'-\mathbf{Z}''$$

where $\mathbf{Z}'\overset{\mathcal{D}}{=}\mathbf{Z}''\overset{\mathcal{D}}{=}\left[\ln X_0^1,\ldots,\ln X_0^k\right]^{\mathsf{T}}$ *and* $\mathbf{Z}'$ *is independent of* $\mathbf{Z}''$*. The hypotheses have to be adequately modified.*
(ii) *This shows that under the stated hypotheses,* $\ln\left(1+\rho_T^i\right)$ *is an estimator of* $\ln\left(1+\rho^i\right)$ *(that in this case is equal to* 0*) such that their difference converges at rate* $O_{\mathbb{P}}\left(T^{-1}\right)$*.*[1] *This rate is quite unusual in Statistical Theory: estimators*

[1]We say that X_n is $O_{\mathbb{P}}(r_n)$ if, for each $\varepsilon>0$, there exists $M>0$ such that:

$$\mathbb{P}\left\{\frac{|X_n|}{r_n}>M\right\}<\varepsilon,\qquad\forall n,$$

where M and ε do not depend on n.

with this feature are called superconsistent. *This is equivalent to the fact that, for any $\varepsilon > 0$, there exists M such that:*

$$\mathbb{P}\left(T \cdot \left|\ln\left(1+\rho_T^i\right) - \ln\left(1+\rho^i\right)\right| > M\right) < \varepsilon,$$
$$\mathbb{P}\left(T \cdot \left|\ln\left(1+\rho_T^i\right)\right| > M\right) < \varepsilon.$$

Indeed, we get, using (A.1) and the basic inequality $\mathbb{P}(X+Y > x+y) \leq \mathbb{P}(X > x) + \mathbb{P}(Y > y)$*:*

$$\begin{aligned}
&\mathbb{P}\left(T \cdot \left|\ln\left(1+\rho_T^i\right)\right| > M\right) = \mathbb{P}\left(\ln\left(\frac{X_T^i}{X_0^i}\right) > M\right) + \mathbb{P}\left(\ln\left(\frac{X_T^i}{X_0^i}\right) < -M\right) \\
&\leq \mathbb{P}\left(\ln X_T^i > M/2\right) + \mathbb{P}\left(\ln X_0^i < -M/2\right) + \mathbb{P}\left(\ln X_T^i < -M/2\right) + \mathbb{P}\left(\ln X_0^i > M/2\right) \\
&= 2 \cdot \mathbb{P}\left\{X_0^i < \exp(-M/2)\right\} + 2 \cdot \mathbb{P}\left\{X_0^i > \exp(M/2)\right\},
\end{aligned}$$

where the final equality derives from the stationarity of $\left(X_t^i\right)_{t\in\mathbb{N}}$.

In the following we will write $\Psi(\lambda) = \mathbb{P}\{Z' - Z'' \leq \lambda\}$ where Z' and Z'' are defined in (2.4). The following Theorem yields a rate for the convergence of $T \cdot \ln\left(1+\rho_T^i\right)$ to its asymptotic distribution: this has the form of a Berry-Esséen bound (see e.g. Breiman, 1992, p. 184), since it bounds the norm between the cumulative distribution function (cdf) of $T \cdot \ln\left(1+\rho_T^i\right)$ and the asymptotic cdf. The interest of this kind of result is that, if we choose $\lambda = \lambda_\alpha$ in such a way that $\Psi(\lambda_\alpha) = (1-\alpha)$, the asymptotic significance level of the test, then the finite sample significance level of the test is compelled to lie in the interval:

$$\alpha + \Delta_T \geq \mathbb{P}\left(T \cdot \ln\left(1+\rho_T^i\right) > \lambda_\alpha\right) \geq \alpha - \Delta_T$$

where $\Delta_T \triangleq \left\|\mathbb{P}\left(T \cdot \ln\left(1+\rho_T^i\right) \leq \lambda\right) - \Psi(\lambda)\right\|_\infty$.

The most interesting property of this Theorem lies in the fact that it links the convergence rate to the mixing coefficient α. This entails that a process with short memory (and more rapidly decreasing α) has a faster convergence rate. In the limiting case of independent and identically distributed observations, the bound is identically equal to 0.

Theorem 2.8. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathbf{X}_t)_{t\in\mathbb{N}}$ a strictly positive process. Suppose, moreover:*
(i) $(\mathbf{X}_t)_{t\in\mathbb{N}}$ is a $2-\alpha-$mixing stationary process with mixing coefficients $\alpha_T \triangleq \alpha\left(\sigma\left(\ln\left(\mathbf{X}_0\right)\right), \sigma\left(\ln\left(\mathbf{X}_T\right)\right)\right)$;
(ii) There exists a positive $p \in (0, \infty]$ and a constant vector $\mathbf{C}$ such that $\left\|\ln\left(\mathbf{X}_0\right) + \mathbf{C}\right\|_p < \infty$;
(iii) $\ln X_t^i$ has, for $i = 1, \ldots, k$ and for any t, a density with respect to the Lebesgue measure that is bounded from above by a strictly positive constant m. Then:

$$\begin{aligned}
&\left\|\mathbb{P}\left\{T \cdot \ln\left(1+\rho_T^i\right) \leq \lambda_i, i = 1, \ldots, k\right\} - \mathbb{P}\left\{\mathbf{Z}' - \mathbf{Z}'' \leq \boldsymbol{\lambda}\right\}\right\|_\infty \\
\leq\ & \frac{(k+2)p+k}{p} \cdot \left(\frac{p \cdot \left(2\sqrt{2}3^{k/2} + 1_{p<\infty}\right)}{2p+k}\right)^{\frac{2p+k}{kp+2p+k}} \\
&\cdot M^{\frac{kp}{kp+2p+k}} \cdot \left\|\ln\left(\mathbf{X}_0\right) + \mathbf{C}\right\|_p^{\frac{kp}{kp+2p+k}} \cdot \alpha_T^{\frac{2p}{kp+2p+k}}
\end{aligned}$$

where:

$$M \geq \max\left\{m, \frac{p\left(2\sqrt{2}3^{k/2}+1_{p<\infty}\right)}{(2p+k)\left\|\ln\left(\mathbf{X}_0\right)+\mathbf{C}\right\|_p}\cdot\alpha_T^{\frac{2p}{2p+k}}\right\}.$$

In the case $p=\infty$:

$$\begin{aligned} &\left\|\mathbb{P}\left\{T\cdot\ln\left(1+\rho_T^i\right)\leq\lambda_i, i=1,\dots,k\right\}-\mathbb{P}\left\{\mathbf{Z}'-\mathbf{Z}''\leq\boldsymbol{\lambda}\right\}\right\|_\infty \\ \leq\quad & (2+k)\cdot M^{\frac{k}{k+2}}\cdot 2^{\frac{1}{k+2}}\cdot 3^{\frac{k}{k+2}}\cdot\left\|\ln\left(\mathbf{X}_0\right)+\mathbf{C}\right\|_\infty^{\frac{k}{k+2}}\cdot\alpha_T^{\frac{2}{k+2}}, \end{aligned}$$

where

$$M\geq\max\left\{m,\frac{\sqrt{2}3^{k/2}}{\left\|\ln\left(\mathbf{X}_0\right)+\mathbf{C}\right\|_\infty}\cdot\alpha_T\right\}.$$

Remark 2.5. (i) *If α_T converges to* 0 *when $T\to\infty$, then for large enough T, M can be taken equal to m.*
(ii) *As an example, let us consider what happens when $p=\infty$ and $\infty>\overline{x}\geq X_0^i\geq\underline{x}>0$ for $i=1,\dots,k$. Then, taking $\mathbf{C}$ composed of elements equal to $-\ln\sqrt{\overline{x}\underline{x}}$, we get $\left\|\ln\left(\mathbf{X}_0\right)+\mathbf{C}\right\|_\infty=\ln\sqrt{\frac{\overline{x}}{\underline{x}}}$ and the Theorem becomes:*

$$\begin{aligned} &\left\|\mathbb{P}\left\{T\cdot\ln\left(1+\rho_T^i\right)\leq\lambda_i, i=1,\dots,k\right\}-\mathbb{P}\left\{\mathbf{Z}'-\mathbf{Z}''\leq\boldsymbol{\lambda}\right\}\right\|_\infty \\ \leq\quad & (2+k)\cdot 2^{\frac{1}{k+2}}\cdot 3^{\frac{k}{k+2}}\cdot\left(M\cdot\ln\sqrt{\frac{\overline{x}}{\underline{x}}}\right)^{\frac{k}{k+2}}\cdot\alpha_T^{\frac{2}{k+2}}. \end{aligned}$$

In the previous paragraphs, we obtained a test statistic and we derived its asymptotic distribution. Here, we show how the asymptotic distribution of $T\cdot\left[\ln\left(1+\rho_T^1\right),\dots,\ln\left(1+\rho_T^k\right)\right]^\mathsf{T}$ can be estimated through kernel estimation. Kernel estimation is a procedure often used to draw inference on a density function (but also on a regression function): consider a sample of N points (say $\mathbf{X}_1,\dots,\mathbf{X}_N$) from the density $f_\mathbf{X}$ on $\mathbb{R}^k$. The *kernel estimator* $\hat{f}_\mathbf{X}^N(\mathbf{x})$ of $f_\mathbf{X}(\mathbf{x})$ is defined as:

$$\hat{f}_\mathbf{X}^N(\mathbf{x})=\frac{1}{Nh_N^k}\sum_{n=1}^N K\left(\frac{\mathbf{x}-\mathbf{X}_n}{h_N}\right)$$

where $K:\mathbb{R}^k\to\mathbb{R}$ is a *kernel*, that is a bounded symmetric density with respect to Lebesgue measure such that:

$$\lim_{\|\mathbf{x}\|\to\infty}\|\mathbf{x}\|^k\cdot K(\mathbf{x})=0,$$

$$\int_{\mathbb{R}^k}\|\mathbf{x}\|^2\cdot K(\mathbf{x})\,\mathrm{d}\mathbf{x}<+\infty.$$

Let $\mathbf{Z}_t=\left[\ln X_t^1,\dots,\ln X_t^k\right]^\mathsf{T}$ and $\mathbf{Z}=\left[\ln X_0^1,\dots,\ln X_0^k\right]^\mathsf{T}$, $f_\mathbf{Z}$ be the probability density function (pdf) of $\mathbf{Z}$ and $f_{-\mathbf{Z}}$ be the pdf of $-\mathbf{Z}$. Let $\hat{f}_\mathbf{Z}^T$ be the estimator of $f_\mathbf{Z}$ based on a sample of T observations, that is:

$$\hat{f}_\mathbf{Z}^T(\mathbf{u})=\frac{1}{Th_T^k}\sum_{t=1}^T K\left(\frac{\mathbf{u}-\mathbf{Z}_t}{h_T}\right). \tag{2.5}$$

We propose to estimate the asymptotic pdf of $T\cdot\left[\ln\left(1+\rho_T^1\right),\ldots,\ln\left(1+\rho_T^k\right)\right]^{\mathsf{T}}$, that is $\left(f_{\mathbf{Z}}*f_{-\mathbf{Z}}\right)(\cdot)$, through $\left(\hat{f}_{\mathbf{Z}}^T*\hat{f}_{-\mathbf{Z}}^T\right)(\cdot)$, where $*$ stands for the convolution operator.

Before stating our result, we propose some notation. Let $\mathcal{C}_{2,k}(b)$ be the class of twice continuously differentiable real-valued functions f, defined on $\mathbb{R}^k$, and such that $\|f\|_\infty\le b$ and $\left\|f^{(2)}\right\|_\infty\le b$ where $f^{(2)}$ is a partial derivative of order 2. Let $\ln^{(m)}x$ be the iterated logarithm of order m, that is the function defined by recursion as $\ln^{(m)}x=\ln\max\left(e,\ln^{(m-1)}x\right)$.

The properties of this estimator are collected in the following result. Remark that we suppose that $f_{\mathbf{Z}}$ is twice continuously differentiable: in some cases of interest $f_{\mathbf{Z}}$ is non-differentiable, but even in this case consistency could be proved under different hypotheses.

Theorem 2.9. *Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space and $(\mathbf{X}_t)_{t\in\mathbb{N}}$ a population process. Moreover, let the following hypotheses hold:*

1. $(\mathbf{X}_t)_{t\in\mathbb{N}}$ *is a stationary* geometrically $\alpha-$mixing *process, that is there exists $c_0>0$ and $\rho\in[0,1)$ such that $\alpha(t)\le c_0\cdot\rho^t$, for $t\ge 1$, where $\alpha(t)$ is defined in Appendix B;*
2. *the density $f_{\mathbf{Z}}$ of $\mathbf{Z}$ with respect to Lebesgue measure belongs to $\mathcal{C}_{2,k}(b)$;*
3. $h_T=c_T\cdot\left(\frac{\ln T}{T}\right)^{\frac{1}{k+4}}$ *and $c_T\to c>0$.*

Then, the asymptotic distribution of $T\cdot\left[\ln\left(1+\rho_T^1\right),\ldots,\ln\left(1+\rho_T^k\right)\right]^{\mathsf{T}}$ as $T\to\infty$ has a density $\left(f_{\mathbf{Z}}*f_{-\mathbf{Z}}\right)$ with respect to Lebesgue measure and a strongly consistent estimator of it is given by $\left(\hat{f}_{\mathbf{Z}}^T*\hat{f}_{-\mathbf{Z}}^T\right)$ where $\hat{f}_{\mathbf{Z}}^T$ is defined by (2.5). Moreover, for any integer m, we have:

$$\frac{1}{\ln^{(m)}T}\cdot\left(\frac{T}{\ln T}\right)^{\frac{2}{k+4}}\left\|\left(f_{\mathbf{Z}}*f_{-\mathbf{Z}}\right)-\left(\hat{f}_{\mathbf{Z}}^T*\hat{f}_{-\mathbf{Z}}^T\right)\right\|_\infty\overset{T\to\infty}{\longrightarrow}0,\qquad\mathbb{P}-\mathsf{as}.$$

Remark 2.6. *The implementation of this estimator can be greatly simplified using some properties of the convolution operator. Indeed, the estimator can be written as:*

$$\begin{aligned}\left(\hat{f}_{\mathbf{Z}}^T*\hat{f}_{-\mathbf{Z}}^T\right)(\mathbf{u})&=\int_{\mathbb{R}^k}\hat{f}_{\mathbf{Z}}^T(\mathbf{u}-\mathbf{v})\cdot\hat{f}_{-\mathbf{Z}}^T(\mathbf{v})\,\mathrm{d}\mathbf{v}\\&=\frac{1}{T^2h_T^{2k}}\sum_{t=1}^T\sum_{s=1}^T\int_{\mathbb{R}^k}K\left(\frac{\mathbf{u}-\mathbf{v}-\mathbf{Z}_t}{h_T}\right)K\left(\frac{\mathbf{v}+\mathbf{Z}_s}{h_T}\right)\mathrm{d}\mathbf{v}\\&=\frac{1}{T^2h_T^k}\sum_{t=1}^T\sum_{s=1}^T(K*K)\left(\frac{\mathbf{u}-\mathbf{Z}_t+\mathbf{Z}_s}{h_T}\right)\end{aligned}$$

that is as a direct kernel estimator. Moreover, using this new formula we have:

$$\begin{aligned}\int_{(-\infty,\lambda]}\left(\hat{f}_{\mathbf{Z}}^{T}*\hat{f}_{-\mathbf{Z}}^{T}\right)(\mathbf{x})\,\mathrm{d}\lambda(\mathbf{x}) &= \frac{1}{T^2h_T^k}\sum_{t=1}^{T}\sum_{s=1}^{T}\int_{(-\infty,\lambda]}(K*K)\left(\frac{\mathbf{x}-\mathbf{Z}_t+\mathbf{Z}_s}{h_T}\right)(\mathbf{x})\,\mathrm{d}\lambda(\mathbf{x})\\ &= \frac{1}{T^2h_T^k}\sum_{t=1}^{T}\sum_{s=1}^{T}\int_{(-\infty,h_T\lambda+\mathbf{Z}_t-\mathbf{Z}_s]}(K*K)(\mathbf{x})\,\mathrm{d}\lambda(\mathbf{x})\\ &= \frac{1}{T^2h_T^k}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathcal{K}(h_T\lambda+\mathbf{Z}_t-\mathbf{Z}_s)\end{aligned}$$

where $\mathcal{K}(\cdot)=\int_{(-\infty,\cdot]}(K*K)(\mathbf{x})\,\mathrm{d}\lambda(\mathbf{x})$.

3 Examples

We show how the previous results can be applied to real data using some series provided by the Time Series Data Library of Rob Hyndman. The series are:

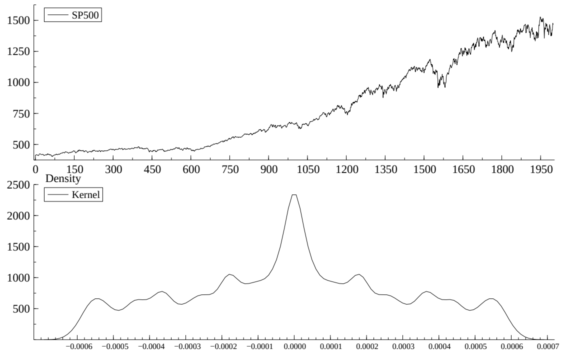
- 2000 daily values of the S&P 500 Index (between July 8, 1992 and June 6, 2000);
- 85 observations of the IPI (Implicit Price Index) in the period 1890-1974, U.S.;
- the quarterly S&P 500 index, in the period 1900-1996;
- the monthly mean water levels in metres of the Lake of the Woods at Warroad - Station No. 05PD001 (at location - latitude 48:54:20N, longitude 095:19:00W) in the period 1916-1965.

For any series, we reproduce in the Table the value of $\bar{R_T}$, $\ln(1+\rho_T)$ and ρ_T, the variance of $\bar{R_T}$ and $\ln(1+\rho_T)$ computed using a Bartlett kernel, the lower and upper bounds of a 95% confidence interval for $\bar{R}$, $\ln(1+\rho)$ and ρ under asymptotic normality. For any series, the graph displays the series and the kernel estimator of the limiting distribution under the hypothesis that $\rho=0$. The peak of the kernel estimator at 0 is due to the fact that T out of T^2 artificial observations used in building it take the value 0. This effect disappears asymptotically.

The value of $\ln(1+\rho_T)$ should be compared with the kernel in order to verify if the series is stationary or not (a value of $\ln(1+\rho_T)$ in the tail witnesses a possible nonstationarity); according to this principle, the first two series are nonstationary, while the last two series seem to be stationary. On the other hand, the intervals around $\ln(1+\rho_T)$ built under the hypothesis of nonstationarity are not able to detect the apparently stationary cases, since they never contain 0.

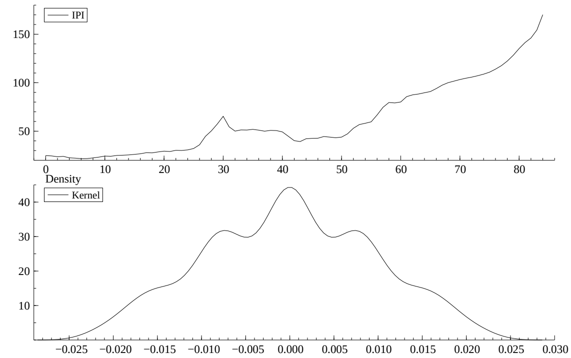
3.1 S&P 500 Index (between July 8, 1992 and June 6, 2000)

	Mean	Variance	Lower bound	Upper bound
$\bar{R}_T$	0.00068176	4.6480e-007	0.00065189	0.00071164
$\ln(1+\rho_T)$	0.00063732	4.0617e-007	0.00060939	0.00066525
ρ_T	0.00063752	-	0.00060957	0.00066547



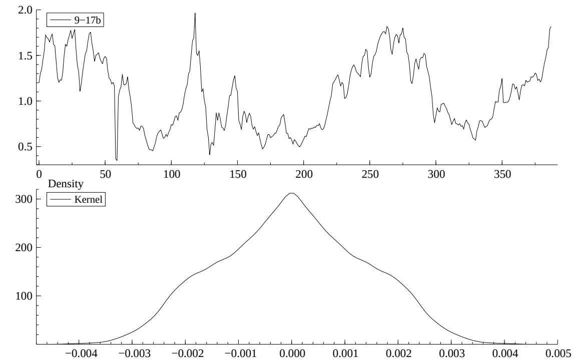
3.2 IPI (Implicit Price Index) in the period 1890-1974, U.S.

	Mean	Variance	Lower bound	Upper bound
$\bar{R}_T$	0.024654	0.00060782	0.019413	0.029895
$\ln(1+\rho_T)$	0.022900	0.00052442	0.018032	0.027769
ρ_T	0.023164	-	0.018195	0.028158



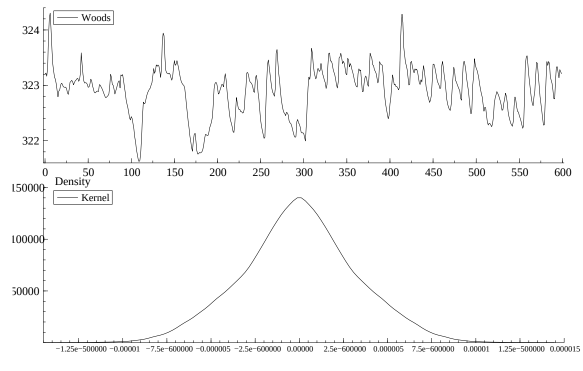
3.3 Quarterly S&P 500 index, 1900-1996

	Mean	Variance	Lower bound	Upper bound
$\bar{R}_T$	0.0075930	5.7653e-005	0.0068374	0.0083485
$\ln(1+\rho_T)$	0.0010750	1.1556e-006	0.00096801	0.0011819
ρ_T	0.0010756	-	0.00096848	0.0011826



3.4 Monthly mean water levels in metres of the Lake of the Woods at Warroad - Station No. 05PD001 (at location - latitude 48:54:20N, longitude 095:19:00W) in the period 1916-1965

	Mean	Variance	Lower bound	Upper bound
$\bar{R}_T$	1.3555e-007	1.8374e-014	1.2470e-007	1.4640e-007
$\ln(1+\rho_T)$	3.6157e-008	1.3073e-015	3.3264e-008	3.9050e-008
ρ_T	3.6157e-008	-	3.3264e-008	3.9050e-008



A Proofs

Proof of Theorem 2.1. The average growth rate is given by (2.1). Using the equation $X_t^i(\omega) = X_0^i(S^t\omega)$, we get:

$$\bar{R}^i(\omega) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \left(\frac{X_1^i(S^t\cdot) - X_0^i(S^t\cdot)}{X_0^i(S^t\cdot)} \right)(\omega).$$

The random variable $\left(\frac{X_1^i(S^t\cdot) - X_0^i(S^t\cdot)}{X_0^i(S^t\cdot)} \right)$ is quasi-integrable (since it is bounded from below by -1, for a definition see Appendix B) and we can apply the extended Ergodic Theorem B.2. We have, $\mathbb{P}-\mathsf{as}$ and $\mathbb{P}^*-\mathsf{as}$:

$$\bar{R}^i(\omega) = \mathbb{E}^* \left[\frac{X_1^i(S^t\cdot) - X_0^i(S^t\cdot)}{X_0^i(S^t\cdot)} |\mathcal{I} \right](\omega),$$

that is:

$$\bar{R}^i(\omega) = \mathbb{E}\left[\frac{X_0^i(S\cdot)}{X_0^i(\cdot)} |\mathcal{I} \right] - 1.$$

The other assertions are trivial.

Proof of Corollary 2.5. Let us proceed by steps, starting from the case in which $\mathbb{P}^*$ is stationary.

1. The first statement is trivial.

2. We have:

$$\bar{R}^i(\omega) = \mathbb{E}\left[\frac{X_0^i(S\cdot)}{X_0^i(\cdot)} |\mathcal{I} \right] - 1 = \mathbb{E}^*\left[\frac{X_0^i(S\cdot)}{X_0^i(\cdot)} |\mathcal{I} \right] - 1$$

and we can use Jensen's inequality to get:

$$\begin{aligned} \bar{R}^i &= \mathbb{E}^*\left[\frac{X_0^i(S\cdot)}{X_0^i(\cdot)} |\mathcal{I} \right] - 1 \\ &= \mathbb{E}^*\left[\exp\left(\ln X_0^i(S\cdot) - \ln X_0^i(\cdot) \right) |\mathcal{I} \right] - 1 \\ &\geq \exp\left\{ \mathbb{E}^*\left[\ln X_0^i(S\cdot) |\mathcal{I} \right] - \mathbb{E}^*\left[\ln X_0^i(\cdot) |\mathcal{I} \right] \right\} - 1 = 0; \end{aligned}$$

moreover, since the exponential is strictly convex, the inequality becomes an equality if and only if $\ln X_0^i(S\cdot) - \ln X_0^i(\cdot)$ is a constant with $\mathbb{P}^*-$probability one. This happens if $R_t^i(\omega) = 0$ for every t but a finite number of constants (that is ultimately).

3. In this case, the statement is a consequence of the previous two.

The case in which $\mathbb{P}^*$ is ergodic can be simply obtained from the previous ones.

Proof of Theorem 2.2. Since $(\mathbf{X}_t)_{t\in\mathbb{N}}$ is ergodic and strongly mixing of constant $\alpha(k)$, also R_t^i will be ergodic and strongly mixing with the same constant. Then we verify the conditions of Theorem B.4 on the random variable $R_t^i(\omega)$; indeed, for $s > 0$:

$$\begin{aligned}
&\mathbb{E}R_t^i = \mu - 1, \\
&\mathbb{E}\left[R_t^i - (\mu - 1)\right]^2 = \sigma^2, \\
&\mathbb{E}\left[\left(R_t^i - \mu + 1\right)\left(R_{t+s}^i - \mu + 1\right)\right] = \sigma^2 \cdot \rho_s.
\end{aligned}$$

Therefore the conditions are satisfied and the Theorem holds.

Proof of Theorem 2.3. Using equation (2.2), the expressions of ρ^i involving the mean with respect to $\mathbb{P}$ and $\mathbb{P}^*$ are a direct consequence of the extended Ergodic Theorem B.2.

Proof of Corollary 2.6. The proof is trivial.

Proof of Theorem 2.4. Using stationarity, we write ρ_T^i as:

$$\ln\left(1+\rho_T^i\right) - \mathbb{E}\ln\left(1+R_t^i\right) = \frac{1}{T}\sum_{t=0}^{T-1}\left[\ln\left(1+R_t^i\right) - \mathbb{E}\ln\left(1+R_t^i\right)\right].$$

Since the conditions for the application of the Central Limit Theorem of Doukhan, Massart and Rio (1994) for strongly mixing processes are satisfied, the result follows. In particular, remark that when Hypothesis **(v)** is not verified, the variance $\sigma^2 \cdot \left(1 + 2 \cdot \sum_{n=1}^{+\infty} \rho_n\right)$ is equal to 0.

Proof of Theorem 2.7. Since $\left(X_t^i\right)_{t=0,1,\dots}$ is stationary and ergodic (since it is mixing, see Appendix B), we have:

$$T \cdot \ln\left(1+\rho_T^i\right) = \ln\left(\frac{X_T^i}{X_0^i}\right),$$

that is:

$$\mathbb{P}\left(T \cdot \ln\left(1+\rho_T^i\right) \le \lambda\right) = \mathbb{P}\left(\ln\left(\frac{X_T^i}{X_0^i}\right) \le \lambda\right). \tag{A.1}$$

Under the hypothesis of $2-\alpha-$mixing stationarity, X_0^i and X_T^i are identically distributed and asymptotically independent as long as $T \to \infty$ and the distribution of $T \cdot \ln\left(1+\rho_T^i\right)$ is as in the statement.

Proof of Theorem 2.8. We start from the following well-known inequality, holding for vectors in $\mathbb{R}^k$ (see e.g. Foata and Fuchs, 1996, p. 213, for the one dimensional case):

$$\Delta \triangleq \left\|\mathbb{P}\{\mathbf{U} \in A\} - \mathbb{P}\{\mathbf{V} \in A\}\right\|_\infty \le \mathbb{P}\left\{\mathbf{V} \in A^{\oplus\varepsilon} \backslash A\right\} + \mathbb{P}\left\{k(\mathbf{U},\mathbf{V}) > \varepsilon\right\} \tag{A.2}$$

where d is a distance in $\mathbb{R}^k$ and $A^{\oplus\varepsilon} \triangleq \left\{\mathbf{u} \in \mathbb{R}^k : \exists \mathbf{v} \in A \text{ such that } d(\mathbf{u},\mathbf{v}) \le \varepsilon\right\}$. We consider the event $\{\mathbf{U} \in A\} = \{\mathbf{U} \le \boldsymbol{\lambda}\} = \{U_i \le \lambda_i, i = 1, \dots, k\}$ for a certain vector $\boldsymbol{\lambda}$ and we take $d(\mathbf{u},\mathbf{v}) = \max_{i=1,\dots,k}|u_i - v_i|$. This allows us to rewrite (A.2) as:

$$\Delta \le \mathbb{P}\left\{\mathbf{V} \in A^{\oplus\varepsilon} \backslash A\right\} + \mathbb{P}\left\{\max_{i=1,\dots,k}|U_i - V_i| > \varepsilon\right\}. \tag{A.3}$$

Now, we take $\mathbf{U} = \left[\ln\left(\frac{X_T^i}{X_0^i}\right)\right]_{i=1,\dots,k}$, $\mathbf{V} = \left[\ln\left(\frac{\widetilde{X}_T^i}{\widetilde{X}_0^i}\right)\right]_{i=1,\dots,k}$ and we take $\ln \widetilde{X}_0^i = \ln X_0^i$ for $i = 1,\dots,k$. We will use the notation $\ln(\mathbf{X}_t)$ to indicate the vector whose $i-$th element is $\ln\left(X_t^i\right)$.

The second term in (A.3) becomes:

$$\mathbb{P}\left\{\max_{i=1,\dots,k}\left|\ln\left(X_T^i\right) - \ln\left(\widetilde{X}_T^i\right)\right| > \varepsilon\right\}$$

and this is amenable to majorization using Thorme 2.1 in Rhomari (2002) (see also Bosq, 1996, Lemma 1.2 and Theorem 3 in Bradley, 1983). The main idea is to start from the pair of (dependent) vectors $(\ln \mathbf{X}_0, \ln \mathbf{X}_T)$ and to build another pair $\left(\ln \widetilde{\mathbf{X}}_0, \ln \widetilde{\mathbf{X}}_T\right)$ of vectors with the same marginal distributions (in particular with $\ln \mathbf{X}_0 = \ln \widetilde{\mathbf{X}}_0$ and $\ln \mathbf{X}_T \overset{\mathcal{D}}{=} \ln \widetilde{\mathbf{X}}_T$, where $\overset{\mathcal{D}}{=}$ means equality in distribution) but such that $\ln \widetilde{\mathbf{X}}_0$ and $\ln \widetilde{\mathbf{X}}_T$ are independent vectors. The Theorem allows not only to build this pair, but also to establish a coupling between $\ln \mathbf{X}_T$ and $\ln \widetilde{\mathbf{X}}_T$, i.e. to build a version of $\ln \widetilde{\mathbf{X}}_T$ that is "near" in distribution to $\ln \mathbf{X}_T$.

In the statement of the Theorem, we let $X \triangleq \ln \widetilde{\mathbf{X}}_0 = \ln \mathbf{X}_0$, $Y \triangleq \ln \mathbf{X}_T$, $Y^* \triangleq \ln \widetilde{\mathbf{X}}_T$, U a uniform random variable (remark that any random variable taking its values in a Borel space can be reduced to a uniform random variable on $[0,1]$) and $\mathbf{C}$ a vector. Then, if $\varepsilon \le \|\ln(\mathbf{X}_T) + \mathbf{C}\|_p$ (we will check this later), we have:

$$\begin{aligned}
&\mathbb{P}\left\{\max_{i=1,\dots,k}\left|\ln\left(X_T^i\right) - \ln\left(\widetilde{X}_T^i\right)\right| > \varepsilon\right\} \\
\le\ & \left(2\sqrt{2}3^{k/2} + 1_{p<\infty}\right) \cdot \left(\frac{\|\ln(\mathbf{X}_0) + \mathbf{C}\|_p}{\varepsilon}\right)^{\frac{kp}{2p+k}} \cdot \alpha_T^{\frac{2p}{2p+k}} \\
=\ & K_{k,p} \cdot \varepsilon^{-\frac{kp}{2p+k}},
\end{aligned}$$

where we have set $\alpha_T \triangleq \alpha(\sigma(\ln(\mathbf{X}_0)), \sigma(\ln(\mathbf{X}_T)))$ and $K_{k,p} \triangleq \left(2\sqrt{2}3^{k/2} + 1_{p<\infty}\right) \cdot \|\ln(\mathbf{X}_0) + \mathbf{C}\|_p^{\frac{kp}{2p+k}} \cdot \alpha_T^{\frac{2p}{2p+k}}$.

The first term in (A.3) can be majorized by:

$$\begin{aligned}
\mathbb{P}\left\{\mathbf{V} \in A^{\oplus\varepsilon}\backslash A\right\} &\le \sum_{i=1}^{k} \mathbb{P}\left\{\lambda_i < V_i \le \lambda_i + \varepsilon\right\} \\
&\le k\varepsilon \cdot \max_{i=1,\dots,k}\max_{x\in(\lambda_i,\lambda_i+\varepsilon]} f_{\ln\left(\widetilde{X}_T^i\right)-\ln\left(\widetilde{X}_0^i\right)}(x),
\end{aligned}$$

and, from the independence of $\widetilde{\mathbf{X}}_0$ and $\widetilde{\mathbf{X}}_T$:

$$f_{\ln\left(\widetilde{X}_T^i\right)-\ln\left(\widetilde{X}_0^i\right)}(x) = \left[f_{\ln\left(\widetilde{X}_T^i\right)} \star f_{-\ln\left(\widetilde{X}_0^i\right)}\right](x) \le \left\|f_{\ln\left(\widetilde{X}_0^i\right)}\right\|_\infty \le m.$$

Therefore, (A.3) becomes:

$$\Delta \le km\varepsilon + K_{k,p} \cdot \varepsilon^{-\frac{kp}{2p+k}}. \qquad \text{(A.4)}$$

The optimal $\varepsilon^{\star}$, obtained differentiating Δ as a function of ε and solving the resulting equation for ε, is given by:

$$\varepsilon^{\star}=\left(\frac{p}{m\left(2p+k\right)}\cdot K_{k,p}\right)^{\frac{2p+k}{kp+2p+k}}$$

and this leads to:

$$\begin{aligned}\Delta^{\star} &\leq \varepsilon\cdot\left(km+K_{k,p}\cdot\varepsilon^{-\frac{2p+k+kp}{2p+k}}\right)\\ &= \frac{\left(k+2\right)pm+km}{p}\cdot\left[\frac{p\cdot K_{k,p}}{m\cdot\left(2p+k\right)}\right]^{\frac{2p+k}{kp+2p+k}}.\end{aligned}$$

In particular, when $p=\infty$ the bound of (A.4) reduces to:

$$\begin{aligned}\Delta &\leq km\varepsilon+2\sqrt{2}3^{k/2}\cdot\left(\frac{\left\|\ln\left(\mathbf{X}_{0}\right)+\mathbf{C}\right\|_{\infty}}{\varepsilon}\right)^{\frac{k}{2}}\cdot\alpha_{T}\\ &= km\varepsilon+K_{k,\infty}\cdot\varepsilon^{-\frac{k}{2}}.\end{aligned}$$

The optimal choice is given by $\varepsilon^{\star}=\left(\frac{K_{k,\infty}}{2m}\right)^{\frac{2}{k+2}}$ and $\Delta^{\star}\leq m\left(2+k\right)\cdot\left(\frac{K_{k,\infty}}{2m}\right)^{\frac{2}{k+2}}$.

Now, we have just to verify that $\varepsilon^{\star}\leq\left\|\ln\left(\mathbf{X}_{0}\right)+\mathbf{C}\right\|_{p}$. Since the bound of (A.4) holds also with m replaced by a larger constant M, we can take M such that:

$$M\geq\max\left\{m,\frac{p\left(2\sqrt{2}3^{k/2}+1_{p<\infty}\right)}{\left(2p+k\right)\left\|\ln\left(\mathbf{X}_{0}\right)+\mathbf{C}\right\|_{p}}\cdot\alpha_{T}^{\frac{2p}{2p+k}}\right\}.$$

Proof of Theorem 2.9. Theorem 2.7 holds. The hypotheses are justified by the use of Lemma 2.1 in Bosq (1996) to show that the estimator is consistent in the topology of uniform convergence. The majorization:

$$\begin{aligned}&\left\|\left(f_{\mathbf{Z}}*f_{-\mathbf{Z}}\right)-\left(\hat{f}_{\mathbf{Z}}^{T}*\hat{f}_{-\mathbf{Z}}^{T}\right)\right\|_{\infty}\\ &=\left\|\left(f_{\mathbf{Z}}*f_{-\mathbf{Z}}\right)-\left(\hat{f}_{\mathbf{Z}}^{T}*f_{-\mathbf{Z}}\right)+\left(\hat{f}_{\mathbf{Z}}^{T}*f_{-\mathbf{Z}}\right)-\left(\hat{f}_{\mathbf{Z}}^{T}*\hat{f}_{-\mathbf{Z}}^{T}\right)\right\|_{\infty}\\ &\leq\left\|\left(f_{\mathbf{Z}}*f_{-\mathbf{Z}}\right)-\left(\hat{f}_{\mathbf{Z}}^{T}*f_{-\mathbf{Z}}\right)\right\|_{\infty}+\left\|\left(\hat{f}_{\mathbf{Z}}^{T}*f_{-\mathbf{Z}}\right)-\left(\hat{f}_{\mathbf{Z}}^{T}*\hat{f}_{-\mathbf{Z}}^{T}\right)\right\|_{\infty}\\ &\leq\left\|f_{\mathbf{Z}}-\hat{f}_{\mathbf{Z}}^{T}\right\|_{\infty}+\left\|f_{-\mathbf{Z}}-\hat{f}_{-\mathbf{Z}}^{T}\right\|_{\infty}=2\cdot\left\|f_{\mathbf{Z}}-\hat{f}_{\mathbf{Z}}^{T}\right\|_{\infty},\end{aligned}$$

proves that $\left\|f_{\mathbf{Z}}-\hat{f}_{\mathbf{Z}}^{T}\right\|_{\infty}\to 0$ is sufficient for our result to hold and yields the desired rate of convergence.

B A Reminder of Stochastic Processes Theory

Consider a stochastic process $(X_t)_{t\in\mathbb{N}}$ defined on a probability space $(\Omega,\mathcal{A},\mathbb{P})$ such that any X_t takes on its values in a measurable space $(\mathfrak{Y},\mathcal{Y})$: for convenience, we consider the coordinate-variable process (see Billingsley, 1979, p.

433), obtained defining $\Omega \triangleq V^{\mathbb{N}}$, $\omega \triangleq (x_t)_{t\in\mathbb{N}}$ and identifying any random variable X_t with the projection operator $X_t(\omega) \triangleq x_t$. Moreover, the process can be represented through the $\mathcal{A}-$measurable shift operator $S : \Omega \to \Omega$ as

$$X_t(\omega) = X_0\left(S^t\omega\right),$$

and the properties of $(X_t)_{t\in\mathbb{N}}$ can be turned into equivalent properties of $\mathbb{P}$.

In the following, we will speak of the transformation S on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and we will often write it as $(\Omega, \mathcal{A}, \mathbb{P}, S)$ or, when no confusion is possible, simply $(\mathbb{P}, S)$. The transformation S is not assumed to be invertible even if S^{-1} is often used: this is due to the fact that $S^{-1}\omega$ is a measurable set and therefore $\mathbb{P}\left(S^{-1}\omega\right)$ is well defined even if $S^{-1}\omega$ is not a single point. Remark that a structure of the form $(\Omega, \mathcal{A}, \mu, S)$, where μ is a measure, is often called a *dynamical system* (see e.g. Lasota and Mackey, 1994).

We define $\mathcal{A}_\infty$ to be the *tail $\sigma-$field*, that is $\bigcap_{n\geq 1}\sigma(X_n, X_{n+1}, ...)$, where $\sigma(Y)$ is the $\sigma-$field generated by the random variable Y.

B.1 Stationarity

An $\mathcal{A}-$measurable transformation $(\Omega, \mathcal{A}, \mathbb{P}, S)$ is said to be *measure-preserving* if $\mathbb{P}\left(S^{-1}A\right) = \mathbb{P}(A)$ for all $A \in \mathcal{A}$: equivalently, $(X_t)_{t\in\mathbb{N}}$ is said to be *stationary*.

B.2 Ergodicity

From now on, we suppose that $(\Omega, \mathcal{A}, \mathbb{P}, S)$ is measure-preserving. The sets $A \in \mathcal{A}$ that satisfy $S^{-1}A = A$ $\mathbb{P}-\mathsf{as}$ (that is $\mathbb{P}\left\{\left(S^{-1}A\right) \triangle A\right\} = \mathbb{P}\left\{\left(S^{-1}A\right)\backslash A\right\} + \mathbb{P}\left\{A\backslash\left(S^{-1}A\right)\right\} = 0$) are called *invariant sets* and constitute a sub$-\sigma-$field $\mathcal{I}$ of $\mathcal{A}$: remark that $\mathcal{I} \subset \mathcal{A}_\infty$ (Proposition 6.32 in Breiman, 1992) while the converse is not true. A random variable X is $\mathcal{I}-$measurable iff $X(\omega) = X(S\omega)$ for $\mathbb{P}-$almost all $\omega \in \Omega$: X is said to be an *invariant random variable*. The $\mathcal{A}-$measurable and measure-preserving transformation $(\Omega, \mathcal{A}, \mathbb{P}, S)$ is said to be *ergodic* or *metrically transitive* if $\mathbb{P}(A) = 0$ or 1 for all invariant sets A; equivalently, $(X_t)_{t\in\mathbb{N}}$ is said to be *ergodic*.

A necessary and sufficient condition for stationary ergodicity is that, for any two sets $A, B \in \mathcal{A}$, the following equation holds:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\mathbb{P}\left(B\cap S^{-j}A\right) = \mathbb{P}(A)\cdot\mathbb{P}(B). \tag{B.1}$$

We can give a characterization of stationary laws in term of mixing of ergodic ones: this is contained in the von Neumann and Kryloff-Bogoliouboff Decomposition Theorem (see Valadier, 2000, for some history), that states that any stationary process can be seen as the result of a mixing operation on the set of ergodic processes.

Theorem B.1. *The set $\mathcal{P}_{\text{stat}}$ of stationary probabilities on $(\Omega, \mathcal{A})$ is a nonempty convex compact subset of the set of probabilities on Ω. The set $\partial\mathcal{P}_{\text{stat}}$ of extreme* *points of $\mathcal{P}_{\text{stat}}$ coincides with the set of ergodic probabilities $\mathcal{P}_{\text{erg}}$. Moreover, if $\mathbb{P} \in \mathcal{P}_{\text{stat}}$, then there is a probability measure λ defined on $\partial\mathcal{P}_{\text{stat}} = \mathcal{P}_{\text{erg}}$ such that:*

$$\forall A \in \mathcal{A}, \qquad \mathbb{P}(A) = \int_{\mathcal{P}_{\text{erg}}} \mathbb{Q}(A)\,\lambda(\mathrm{d}\mathbb{Q}).$$

Remark that, especially in Markov chains, alternative concepts of ergodicity that do not need stationarity can be encountered. However, since Markov chains are not used in this paper, we will avoid introducing these concepts.

B.3 Asymptotic Mean Stationarity

An $\mathcal{A}-$measurable transformation S on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is said to be *asymptotically mean stationary* (see Gray and Kieffer, 1980, for definitions and properties) if $(1/n)\sum_{j=0}^{n-1}\mathbb{P}\left(S^{-j}A\right)$ is convergent for all $A \in \mathcal{A}$. A probability measure $\mathbb{P}'$ is said to *asymptotically dominate* $\mathbb{P}$ if $\mathbb{P}'(A) = 0$ for $A \in \mathcal{A}_\infty$ implies that $\lim_{n\to\infty}\mathbb{P}(S^{-n}A) = 0$.

The following useful facts show some properties of asymptotic mean stationarity and of the probability $\mathbb{P}^*(\cdot) = \lim_{n\to\infty}(1/n)\sum_{j=0}^{n-1}\mathbb{P}\left(S^{-j}\cdot\right)$ (see Barron, 1985, pp. 1298-1299, for the first two statements, and Gray and Kieffer, 1980, p. 964, for the third one):

- if $\mathbb{P}$ is asymptotically mean stationary, then $\mathbb{P}^*$ is a stationary measure that asymptotically dominates $\mathbb{P}$;
- if $\mathbb{P}'$ is stationary, then asymptotic dominance is equivalent to absolute continuity of the measures restricted to the tail $\sigma-$field: if $A \in \mathcal{A}_\infty$ and $\mathbb{P}'(A) = 0 \Rightarrow \mathbb{P}(A) = 0$;
- $\mathbb{P}$ and $\mathbb{P}^*$ have the same restriction to the invariant $\sigma-$field $\mathcal{I}$.

In the following, we state a version of the Ergodic Theorem for quasi-integrable extended-real valued random variables as proved in Choirat *et al.* (2005). We recall that a random variable Y is said to be *quasi-integrable* in the sense of Neveu (1964, p. 40) if it satisfies the hypothesis that either $\mathbb{E}\max\{Y, 0\}$ or $\mathbb{E}\min\{Y, 0\}$ is finite.

Theorem B.2. *Let $(\Omega, \mathcal{A}, \mathbb{P}, S)$ be an $\mathcal{A}-$measurable asymptotically mean stationary transformation. Then, for every quasi-integrable extended real-valued positive random variable Y, one has*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}Y\left(S^i\omega\right) = \mathbb{E}^*\left(Y\,|\mathcal{I}\right)(\omega), \qquad \mathbb{P}-\mathsf{as}\ \textit{and}\ \mathbb{P}^*-\mathsf{as},$$

(where both sides can be equal to $+\infty$ or $-\infty$).

B.4 Mixing

As shown by equation (B.1), ergodicity involves a requirement on the Cesaro mean of the probability; on the other hand, mixing involves the same behavior on the probability, in the sense that for any A and B:

$$\lim_{n\to\infty} \mathbb{P}\left(B \cap S^{-n}A\right) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Let $\mathcal{B}$ and $\mathcal{C}$ be sub$-\sigma-$fields of $\mathcal{A}$; define:

$$\alpha(\mathcal{B},\mathcal{C}) \triangleq \sup_{B\in\mathcal{B}, C\in\mathcal{C}} |\mathbb{P}(B\cap C) - \mathbb{P}(B)\cdot\mathbb{P}(C)|,$$

$$\phi(\mathcal{B},\mathcal{C}) \triangleq \sup_{B\in\mathcal{B}, C\in\mathcal{C}, \mathbb{P}(B)>0} |\mathbb{P}(C|B) - \mathbb{P}(C)|.$$

Now, let $\mathcal{A}_s^t = \sigma(Y_t, \ldots, Y_s)$ where $s \geq t$ and t can be equal to $-\infty$ and s to $+\infty$. The transformation $(\Omega, \mathcal{A}, \mathbb{P}, S)$ is *strong* (or $\alpha-$) *mixing* if:

$$\alpha(n) \triangleq \sup_{k\in\mathbb{Z}} \alpha\left(\mathcal{A}_{-\infty}^k, \mathcal{A}_{k+n}^{+\infty}\right) \overset{n\to\infty}{\longrightarrow} 0;$$

it is *uniform* (or $\phi-$) *mixing* if:

$$\phi(n) \triangleq \sup_{k\in\mathbb{Z}} \phi\left(\mathcal{A}_{-\infty}^k, \mathcal{A}_{k+n}^{+\infty}\right) \overset{n\to\infty}{\longrightarrow} 0.$$

Another concept is $2-\alpha-$*mixing* (see Bosq, 1996, p. 17). The time series $(X_t)_{t\in\mathbb{Z}}$ on $(\Omega, \mathcal{A}, \mathbb{P})$ is $2-\alpha-$*mixing* if:

$$\alpha^{(2)}(n) \triangleq \sup_{k\in\mathbb{Z}} \alpha(\sigma(X_k), \sigma(X_{k+n})) \overset{n\to\infty}{\longrightarrow} 0.$$

The following Lemma will be used in the proof of one of our main results.

Lemma B.3 (Bradley). *Let (X,Y) be a $\mathbb{R}^d \times \mathbb{R}-$valued random vector such that $Y \in L^p(\mathbb{P})$ for some $p \in [1, +\infty]$. Let c be a real number such that $\|Y+c\|_p > 0$, and $\xi \in \left(0, \|Y+c\|_p\right]$. Then, there exists a random variable Y^* such that*

1. *Y^* and Y have the same distribution and Y^* is independent of X,*
2. $\mathbb{P}(|Y^* - Y| > \xi) \leq 11 \cdot \left(\xi^{-1}\|Y+c\|_p\right)^{p/(2p+1)} \cdot [\alpha(\sigma(X), \sigma(Y))]^{2p/(2p+1)}.$

It is simple to see that:

$$\begin{aligned}
\alpha(n) &= \sup_{k\in\mathbb{Z}} \sup_{B\in\mathcal{A}_{-\infty}^k, C\in\mathcal{A}_{k+n}^{+\infty}} |\mathbb{P}(B\cap C) - \mathbb{P}(B)\cdot\mathbb{P}(C)| \\
&= \sup_{k\in\mathbb{Z}} \sup_{B\in\mathcal{A}_{-\infty}^k, C\in\mathcal{A}_{k+n}^{+\infty}} |\mathbb{P}(C|B) - \mathbb{P}(C)| \cdot \mathbb{P}(B) \\
&\leq \left\{ \sup_{k\in\mathbb{Z}} \sup_{B\in\mathcal{A}_{-\infty}^k, C\in\mathcal{A}_{k+n}^{+\infty}} |\mathbb{P}(C|B) - \mathbb{P}(C)| \right\} \cdot \left\{ \sup_{k\in\mathbb{Z}} \sup_{B\in\mathcal{A}_{-\infty}^k} \mathbb{P}(B) \right\} \\
&\leq \sup_{k\in\mathbb{Z}} \sup_{B\in\mathcal{A}_{-\infty}^k, C\in\mathcal{A}_{k+n}^{+\infty}, \mathbb{P}(B)>0} |\mathbb{P}(C|B) - \mathbb{P}(C)| = \phi(n),
\end{aligned}$$

that is:

$$\lim_{n\to\infty} \phi(n) = 0 \Rightarrow \lim_{n\to\infty} \alpha(n) = 0,$$

and the uniform mixing concept is stronger than the strong mixing one. Moreover, $2-\alpha-$mixing is even weaker than $\alpha-$mixing (Bosq, 1996).

The following Central Limit Theorem holds for strongly mixing processes (see Doukhan *et al.*, 1994).

Theorem B.4 (Doukhan-Massart-Rio). *Let $(X_n)_{n\in\mathbb{N}}$ be a stationary and strong mixing process with $\mathbb{E}X_n = 0$, $\mathbb{E}X_n^2 = \bar{\sigma}^2 \in]0,\infty[$, $\mathbb{E}X_{n+s}X_n = \bar{\sigma}^2\bar{\gamma}_s$ and suppose that for some real number $r>2$, $\mathbb{E}|X_n|^r < \infty$ and $\sum_{n=1}^{+\infty} n^{\frac{2}{r-2}}\alpha(n) < +\infty$. Then $\sum_{n=1}^{+\infty} |\bar{\gamma}_n| < +\infty$ and the following limit holds*

$$\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} X_n \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \bar{\sigma}^2 \cdot \left(1 + 2\cdot\sum_{n=1}^{+\infty}\bar{\gamma}_n\right)\right).$$

References

[1] ANDREWS, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**, 817-858.

[2] ANDREWS, D.W.K. and MONAHAN, J.C. (1992). An improved heteroscedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica* **60**, 953-966.

[3] BARRON, A.R. (1985), The strong ergodic theorem for densities: Generalized Shannon-McMillan-Breiman theorem, *Annals of Probability*, **13**, pp. 1292-1303.

[4] BILLINGSLEY, P. (1979). *Probability and Measure*, Wiley, New York.

[5] BOSQ, D. (1996). *Nonparametric Statistics for Stochastic Processes*, Lecture Notes in Statist., Vol. 110, Springer-Verlag, New York.

[6] BRADLEY, R.C. (1983). Approximating theorems for strongly mixing random variables, *Michigan Math. J.* **30**, 69-81.

[7] BREIMAN, L. (1992). *Probability*, 2nd edition, SIAM, Philadelphia.

[8] CHOIRAT, C., HESS, C. and SERI, R. (2003). A functional version of the Birkhoff ergodic theorem for a normal integrand: A variational approach, *The Annals of Probability*.

[9] CHOIRAT, C., HESS, C. and SERI, R. (2005). An extended Birkhoff ergodic theorem, Working Paper.

[10] DAVIDSON, J. (2000). *Econometric Theory*, Blackwell Publishers.

[11] HUANG, W. and DAY, R.H. (2001). On the Statistical Properties of Ergodic Economic Systems, *Discrete Dynamics in Nature and Society* **6**, 181-189.

[12] DE JONG, R.M. and DAVIDSON, J. (2000). Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices, *Econometrica* forthcoming.

[13] DOUKHAN, P., MASSART, P. and RIO, E. (1994). The functional central limit theorem for strongly mixing processes, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **10**, 63-82.

[14] FOATA, D. and FUCHS, A. (1996). *Calcul des Probabilités*, Masson, Paris.

[15] GRAY, R.M. and KIEFFER, J.C. (1980). Asymptotically mean stationary measures, *Ann. Probab.* **8**, 962-973.

[16] HAMILTON, J.D. (1994). *Time Series Analysis*, Princeton University Press.

[17] Huang, W. (2004). On the long-run average growth rate of chaotic systems, *Chaos* **14**, 38-47.

[18] Katok, A. and Hasselblatt, B. (1995). *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and its Applications, Vol. 54, Cambridge University Press.

[19] Lasota, A. and Mackey, M.C. (1994). *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, 2nd edition, Springer-Verlag, New York.

[20] Liverani, C. (1996), Central Limit Theorem for deterministic systems, *Working Paper*.

[21] Newey, W.K. and West, K.D. (1987). A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703-708.

[22] Newey, W.K. and West, K.D. (1994). Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* **61**, 631-654.

[23] Neveu, J. (1964). *Bases Mathématiques du Calcul des Probabilités*, Masson, Paris.

[24] Rhomari, N. (2002). Approximation et inégalités exponentielles pour les sommes de vecteurs aléatoires dépendants, *C. R. Acad. Sci. Paris, Ser. I* **334**, 149154.

[25] Rio, E. (1995). About the Lindeberg method for strongly mixing sequences, *ESAIM: Probability and Statistics* **1**, 35-61.

[26] Valadier, M. (2000). Survey on some questions about stationary and ergodic processes, *Working Paper*, Université de Montpellier.