Econometric Theory, **15**, 1999, 643–663. Printed in the United States of America.

DECIDING BETWEEN *I*(0) AND *I*(1) VIA FLIL-BASED BOUNDS

VALENTINA CORRADI
University of London

We construct properly scaled functions of R^p-valued partial sums of demeaned data and derive bounds via the functional law of the iterated logarithm for strong mixing processes. If we obtain a value below or equal to the bound we decide in favor of $I(0)$; otherwise we decide in favor of $I(1)$. This provides a consistent rule for classifying time series as being $I(1)$ or $I(0)$. The nice feature of the procedure lies in the almost sure nature of the bound, guaranteeing a lim sup–type result. We finally provide conditions for the strong consistency of estimators of the variance in the dependent and heterogeneous case.

1. INTRODUCTION

Consistent rules for deciding whether a time series is integrated of order one, $I(1)$, or is integrated of order zero, $I(0)$, have been recently proposed (see, e.g., Phillips and Ploberger 1994, 1996; Stock 1994). The main idea underlying these papers is to construct a Bayes factor, or a posterior odd ratio, such that, under the $I(0)$,($I(1)$) hypothesis the ratio converges in probability to zero and under the $I(1)$,($I(0)$) hypothesis diverges to infinity. According to Phillips and Ploberger's PIC (posterior information criterion), we choose the $I(1)$ hypothesis versus the $I(0)$ if the Bayes model likelihood ratio is less than one and vice versa. On the other hand, Stock (1994) does not study the posterior distribution of the data directly, but he constructs the likelihood ratio of observing some functionals of partial sums of the (demeaned or detrended) data under either the $I(0)$ or the $I(1)$ hypothesis; the classification rates are then computed by direct simulation. More recently Sin (1996), extending previous work by Sin and White (1996), proposed a flexible information criterion for selecting strictly nested, linear, possibly nonstationary models. According to Sin's information criterion, we choose unit root if we obtain a negative value, and we choose no unit root if we obtain a strictly positive value; this also delivers a completely consistent procedure.

In this paper we propose a new completely consistent rule for deciding between $I(1)$ and $I(0)$, from a "non-Bayesian" point of view. The main feature of the

I thank a co-editor, Bruce Hansen, three anonymous referees, Filippo Altissimo, Frank Diebold, Jin Hahn, Atsushi Inoue, Lutz Kilian, Serena Ng, Shinichi Sakata, Chor-Yiu Sin, and the seminar participants at the University of Pennsylvania, University of Florida, Econometric Society 1996 Winter Meeting for very helpful comments and suggestions. Address correspondence to: Valentina Corradi, Department of Economics, Queen Mary and Westfield College, University of London, Mile End Road, London, E1 4NS, UK; e-mail: v.corradi@qmw.ac.uk.

© 1999 Cambridge University Press 0266-4666/99 $9.50

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

procedure we shall consider lies in the almost sure nature of the bound. In what follows we define a purely stochastic process to be $I(d)$ if its dth differences are strong mixing and satisfy some mild moment conditions. Here we concentrate on the demeaned case. The definition of an $I(d)$ process as a process whose dth differences are strong mixing is slightly more restrictive than the more common definition of an $I(d)$ process as a process whose dth differences have partial sums that, properly scaled, satisfy a functional central limit theorem.

We construct a properly scaled functional of the Euclidean norm of R^p-valued partial sums of demeaned data. Under mild moment conditions, partial sums of R^p-valued strong mixing processes satisfy an almost sure invariance principle (e.g., Eberlain 1986). It then follows that the set of norm limit points of such partial sums coincides, almost surely, with the set of norm limit points of the p-dimensional Brownian motion. We then obtain a functional law of the iterated logarithm (FLIL) for the Euclidean norm of the partial sums of R^p-valued strong mixing processes. Finally, using a result by Strassen (1964) that plays, in an almost sure context, the same role that the continuous mapping theorem plays in a weak convergence context, we obtain an almost sure bound for functionals of the Euclidean norm of partial sums of strong mixing processes. More precisely, we shall consider the following statistic: $V_T = b_T^{-1}\sum_{t=1}^{T}\|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j - \bar{X})\|$, where $b_T = T^{3/2}\sqrt{2\log\log T}$, log denotes the natural logarithm, X_t is a p-dimensional process, $\hat{\Omega}_T$ is a proper estimator for the variance, and $\bar{X}$ denotes the sample mean. We shall show that, if X_t is a strong mixing process, then the lim sup of V_T is, almost surely, less than or equal to $1/\sqrt{3} + \frac{1}{2}$ for the case of $p = 1$, whereas for $p > 1$, the lim sup of V_T is almost surely less than or equal to $\frac{2}{3} + \frac{1}{2}$. On the other hand, if X_t is an integrated process, or if at least one component of X_t is $I(1)$, then V_T diverges with probability one at rate $\sqrt{T}/\sqrt{2l_T \log\log(T)}$, where l_T is the lag truncation parameter used in the estimation of the variance. Thus we decide in favor of $I(0)$ whenever we get a value for the statistic smaller than or equal to $1/\sqrt{3} + \frac{1}{2}$, for $p = 1$, or smaller than or equal to $\frac{2}{3} + \frac{1}{2}$, for $p > 1$, and we decide instead in favor of $I(1)$ whenever we get a larger value.

The statistic proposed here is very similar to the statistic proposed by Kwiatkowski, Phillips, Schmidt, and Shin (1992). In principle we could have used their statistic and let the critical values depend on the sample size. Although several choices of sequences of critical values assure the zero size property, it is not clear which sequence would also ensure the almost sure nature of the bound.[1]

To preserve the almost sure nature of the bound we need $\hat{\Omega}_T$ to converge almost surely to Ω_0. We then provide a proof of strong consistency for Newey–West–type variance estimators.

The paper is organized as follows. Section 2 states the assumptions and some preliminary results. Section 3 provides a proof of strong consistency for Newey–West–type variance estimators. Finally Section 4 provides a consistent rule for choosing between $I(0)$ and $I(1)$. All the proofs are collected in the Appendix.

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

2. ASSUMPTIONS AND PRELIMINARIES

We consider the following data generating process (DGP), which is a multivariate extension of the DGP considered by Kwiatkowski et al. (1992):

$$X_t = \mu + \epsilon_t + r_t, \tag{1}$$

$$r_t = r_{t-1} + \xi_t, \tag{2}$$

where X_t, ϵ_t, r_t, and ξ_t, are R^p-valued random variables and $\xi_t \sim$ i.i.d.$(0,\Sigma)$. We can formulate the $I(0)$ hypothesis as

H_0: Σ is a $p \times p$ matrix of zeros,

and we can formulate the $I(1)$ hypothesis as

H_A: Σ is non-negative definite, and there is an $r \times r$ block Σ_r, $1 \leq r \leq p$, which is positive definite.

Thus the $I(1)$ hypothesis, H_A, is meant as the case in which at least one component of X_t is integrated.

In the sequel we shall need the following assumptions.

A1. $\{\epsilon_t\}_{t=-\infty}^{\infty}$ is a zero mean R^p-valued process, such that $\sup_t E(|\epsilon_{it}|^{2(4+\delta)}) \leq C_1 < \infty$, $\forall i = 1,\ldots p$, where $\delta > 0$, and ϵ_{it} denotes the ith component of ϵ_t.

A2. $\{\epsilon_t\}_{t=-\infty}^{\infty}$ is α-mixing with size $-(4(4+\delta))/\delta$, where δ is the same positive constant defined in A1.

Let $S_T(m) = \sum_{k=m+1}^{m+T} \epsilon_k$ and let $S_{iT}(m)$ be the ith component of $S_T(m)$, hereafter $S_T = S_T(0)$.

A3. There exists a positive definite matrix $\Omega_0 = [\omega_{0,ij}]$, such that $\forall i,j = 1,2,\ldots p$,

$$|T^{-1}E(S_{iT}(m)S_{jT}(m)) - \omega_{0,ij}| \leq C_2 T^{-\psi},$$

where $\psi > 0$ and C_2 is independent of m.

Remarks.

(1) Assumptions A1 and A2 are standard memory and moment conditions; as usual we note a trade-off between the degree of heterogeneity and the degree of dependence we can allow for. Assumption A3 is a condition on the rate of growth of the variance of the partial sums of ϵ_t, which, as we shall see subsequently, can be relaxed in the covariance stationary case.

(2) Under the $I(0)$ hypothesis X_t is a strong mixing process, around a possibly nonzero mean. It is worthwhile to notice that, as pointed out by Andrews (1984), there are examples of $I(0)$ processes that are not strong mixing.

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

In the sequel we shall consider almost sure bounds, under both the $I(0)$ and the $I(1)$ hypothesis, for the following statistic:

$$V_T = \frac{1}{T^{3/2}\sqrt{2 \log\log T}} \sum_{t=1}^{T} \left\| \hat{\Omega}_T^{-1/2} \sum_{j=1}^{T} (X_j - \bar{X}) \right\|, \tag{3}$$

where $\bar{X} = T^{-1} \sum_{t=1}^{T} X_t$, $\|\cdot\|$ denotes the Euclidean norm, and $\hat{\Omega}_T$ is a heteroskedasticity-autocorrelation robust estimator of $\mathrm{Var}(T^{-1/2} \sum_{t=1}^{T} X_t)$, as defined in (6), which follows. To derive an almost sure bound for V_T, we need the following preliminary result.

LEMMA 2.1. *Let* A1–A3 *hold. Then there exists a p-dimensional Brownian motion B with covariance matrix equal to* Ω_0, *such that, as* $T \to \infty$,

(i)

$$\sup_{r \in [0,1]} \frac{1}{\sqrt{2T \log\log T}} \left\| \Omega_0^{-1/2} \sum_{k=1}^{[\mathrm{Tr}]} \epsilon_k - W(\mathrm{Tr}) \right\| \xrightarrow{a.s.} 0,$$

(ii) *almost surely*,

$$\limsup_{T \to \infty} \left\| \frac{\Omega_0^{-1/2}}{\sqrt{2T \log\log T}} \sum_{k=1}^{T} \epsilon_k \right\| = 1.$$

In the covariance stationary case we can relax A3. In fact, we have the following corollary.

COROLLARY 2.2. *Let* ϵ_t *be a covariance stationary process satisfying* A1 *and* A2 *and such that* $\forall i = 1, \ldots p$ $\mathrm{Var}(S_{iT}) \geq r(T)$, *with* $r(T) \to \infty$ *as* $T \to \infty$. *Then* ϵ_t *satisfies* A3.

Lemma 2.1(i) provides a strong invariance principle for strong mixing processes satisfying A1–A3 and so for weakly stationary strong mixing processes satisfying A1 and A2. Strong invariance principles ensure that the set of norm limit points of properly scaled partial sums coincides, with probability one, with the set of norm limit points of the Brownian motion. This result will be used to derive an almost sure bound for the statistic V_T. Part (ii) of Lemma 2.1 is just the ordinary law of the iterated logarithm for strong mixing processes, and, as shown in the proof, it can be derived as a corollary of part (i).

3. A STRONG CONSISTENT ESTIMATOR FOR THE LONG-RUN VARIANCE

To have an almost sure bound free of nuisance parameters, we need to rescale the partial sums by the variance; as in general we do not know the variance, we need to estimate it. Although for inference purposes, it suffices to have a weakly consistent estimator, in this context, to preserve the almost sure nature of the bound,

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

we need a strongly consistent estimator. Weak consistency results for estimators of the variance in the heterogeneous and dependent case have been already obtained, in different frameworks, by several authors, among others White (1984), Newey and West (1987, 1994), Gallant and White (1988), Robinson (1991), Andrews (1991), Andrews and Monahan (1992), and Hansen (1992). However, to the best of our knowledge, no strong consistency result has been provided.

Let $\hat{\Omega}_T$ be an estimator for Ω_0 and for notational simplicity assume that $\hat{\Omega}_T^{-1/2}$ is well defined (we could otherwise use a generalized inverse). We then have

$$\begin{aligned}
&\frac{1}{\sqrt{2T \log\log T}} \left\| \hat{\Omega}_T^{-1/2} \sum_{k=1}^{[\mathrm{Tr}]} \epsilon_k \right\| \\
&\quad = \frac{1}{\sqrt{2T \log\log T}} \left\| \Omega_0^{-1/2} \sum_{k=1}^{[\mathrm{Tr}]} \epsilon_k + (\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2}) \sum_{k=1}^{[\mathrm{Tr}]} \epsilon_k \right\| \\
&\quad = \frac{1}{\sqrt{2T \log\log T}} \left\| \Omega_0^{-1/2} \sum_{k=1}^{[\mathrm{Tr}]} \epsilon_k \right\| + o_{as}(1)
\end{aligned}$$

if $\hat{\Omega}_T - \Omega_0 = o_{as}(1)$.

Let $S_T = \sum_{k=1}^{T} \epsilon_k$ and

$$\Omega_0 = \lim_{T\to\infty} E\left(\frac{1}{T} S_T S_{T'}\right),$$

$$\bar{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T} E(\epsilon_t \epsilon_{t'}) + \frac{1}{T}\sum_{\tau=1}^{T-1}\sum_{t=\tau+1}^{T} E(\epsilon_t \epsilon_{t-\tau'} + \epsilon_{t-\tau}\epsilon_{t'}),$$

$$\tilde{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T} E(\epsilon_t \epsilon_{t'}) + \frac{1}{T}\sum_{\tau=1}^{l_T}\sum_{t=\tau+1}^{T} E(\epsilon_t \epsilon_{t-\tau'} + \epsilon_{t-\tau}\epsilon_{t'}). \tag{4}$$

Let $w_\tau = (1 - (\tau/l_T + 1))$ and let

$$\ddot{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \epsilon_{t'} + \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} (\epsilon_t \epsilon_{t-\tau'} + \epsilon_{t-\tau}\epsilon_{t'}), \tag{5}$$

$$\begin{aligned}
\hat{\Omega}_T = &\frac{1}{T}\sum_{t=1}^{T} (X_t - \bar{X})(X_t - \bar{X})' \\
&+ \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} ((X_t - \bar{X})(X_{t-\tau} - \bar{X})' + (X_{t-\tau} - \bar{X})(X_t - \bar{X})').
\end{aligned} \tag{6}$$

To prove the strong consistency of $\ddot{\Omega}_T$ for Ω_0, we show that

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} (\epsilon_t \epsilon_{t-\tau'} - E(\epsilon_t \epsilon_{t-\tau'})) = o_{as}(1). \tag{7}$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

This will be accomplished in Lemma 3.1, which follows. Then it will be easy to show that $\ddot{\Omega}_T - \hat{\Omega}_T = o_{as}(1)$.

It is immediate to see that in the martingale difference sequences (m.d.s.) case, a consistent estimator for the variance is given by $T^{-1}\sum_{t=1}^{T}(X_t - \bar{X})(X_t - \bar{X})'$; thus the desired result follows straightforwardly from the strong law of large numbers. On the other hand, in the non-m.d.s. case we can no longer appeal to the strong law of large numbers.

LEMMA 3.1. *Let* A1 *and* A2 *hold. If as* $T \to \infty$, $l_T \to \infty$ *and* $l_T/T^{1/8} \to 0$, *then*

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=1}^{T-\tau}(\epsilon_t \epsilon_{t+\tau'} - E(\epsilon_t \epsilon_{t+\tau})) = o_{as}(1).$$

We are now ready to show the strong consistency of $\hat{\Omega}_T$ for Ω_0, under the $I(0)$ hypothesis, i.e., as if H_0 were true. The limiting behavior of $\hat{\Omega}_T$ under the $I(1)$ hypothesis will be considered in the next section.

THEOREM 3.2. *Let* A1–A3 *hold. If as* $T \to \infty$, $l_T \to \infty$ *and* $l_T/T^{1/8} \to 0$, *then under* H_0,

$$\hat{\Omega}_T - \Omega_0 = o_{as}(1)$$

To obtain a strong consistency result for the estimator of the long-run variance, we impose a rate of growth for the lag truncation parameter much slower than the rate allowed in the weak consistency case. However it is possible to allow for a faster (but slower than $T^{1/4}$) rate of growth of l_T, if we are willing to strengthen the moment and the memory conditions stated in A1 and in A2. We have the following corollary.

COROLLARY 3.3. *Let* ϵ_t *be an* R^p-*valued zero-mean process such that for some* $\beta \geq 2$ *and* $\delta > 0$, $\sup_t E(\epsilon_{it}^{2(2\beta+\delta)}) \leq C_1 < \infty$ *for all* $i \leq p$ *and* ϵ_t *is* α*-mixing with size* $-2\beta(2\beta+\delta)/\delta$. *If* ϵ_t *satisfies* A3, *and* $l_T \to \infty$ *such that* $l_T = o(T^{(\beta-1)/4\beta})$, *then the statements in Lemma* 3.1 *and in Theorem* 3.2 *hold.*

4. AN ALMOST SURE RULE FOR DECIDING BETWEEN *I*(0) AND *I*(1)

In this section we provide an almost sure rule for deciding between $I(0)$ and $I(1)$, i.e., for deciding in favor of H_0 or in favor of H_A. It will be shown that V_T, as defined in (3), under H_0, as $T \to \infty$ is almost surely smaller than or equal to a certain number, whereas under H_A it diverges with probability one.

THEOREM 4.1. *Let* X_t *be an* R^p*-valued process generated as in* (1) *and* (2) *and let* ϵ_t *satisfy* A1–A3. *If, as* $T \to \infty$, $l_T \to \infty$ *and* $l_T/T^{1/8} \to 0$, *then under* H_0, *almost surely*

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

(i) *for* $p = 1$,

$$\limsup_{T\to\infty} V_T \leq \frac{1}{\sqrt{3}} + \frac{1}{2}$$

(ii) *for* $p > 1$,

$$\limsup_{T\to\infty} V_T \leq \frac{2}{3} + \frac{1}{2}$$

where V_T *is defined as in* (3).

From the preceding theorem we see that for the multidimensional case, regardless of the specific value of p, we obtain a slightly looser almost sure bound. In fact, for $p = 1$ we know from Strassen (1964, p. 219) that

$$\limsup_{T\to\infty} \frac{1}{T^{3/2}\sqrt{2\log\log T}} \sum_{t=1}^{T} \left\| \Omega_0^{-1/2} \sum_{j=1}^{t} \epsilon_j \right\| = \frac{1}{\sqrt{3}}.$$

However, the proof of such a result is taylored for the scalar case and cannot be generalized to the vector case. For the case of $p > 1$, we show in the proof of Theorem 4.1 that

$$\limsup_{T\to\infty} \frac{1}{T^{3/2}\sqrt{2\log\log T}} \sum_{t=1}^{T} \left\| \Omega_0^{-1/2} \sum_{j=1}^{t} \epsilon_j \right\| \leq \frac{2}{3}.$$

Thus for $p > 1$, regardless of the specific value of p, we obtain an almost sure upper bound rather than an exact almost sure bound. The difference between the two bounds is equal to 0.089. Finally for all finite $p = 1, 2, \ldots$ the additional $\frac{1}{2}$ term appearing in (i) and (ii) of Theorem 4.1 is due to the fact that X_t has an unknown possibly nonzero mean and so we construct the statistic using deviations from the sample mean.

In the proof of the preceding theorem we make extensive use of results by Strassen (1964). Now Strassen considers the case of identically and independently distributed random variables, whereas we consider the case of heterogeneous and dependent observations. However, this does not create any problem: in fact we use a strong invariance principle for dependent and heterogeneous observations (e.g., Eberlain, 1986, Theorem 1) to prove the statement in Lemma 2.1(i). Then we can rely on the results by Strassen (1964) to find the lim sup of $(1/\sqrt{2T\log\log T})W(T\cdot)$ and of its functionals.

We now analyze the behavior of V_T under H_A, i.e., when at least one component of X_t is an integrated process.

THEOREM 4.2. *Let* X_t *be an* R^p*-valued process generated as in* (1) *and* (2), *and let* A1–A3 *hold. If as* $T \to \infty$, $l_T \to \infty$ *and* $l_T/T^{1/8} \to 0$, *then under* H_A,

$$P\left(\lim_{T\to\infty} \frac{\sqrt{2l_T\log\log T}}{\sqrt{T}} V_T > 0 \right) = 1.$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

From Theorem 4.2 we see that under the $I(1)$ hypothesis V_T diverges with probability one at rate $\sqrt{T/2l_T \log\log(T)}$. Thus to decide in favor of the $I(0)$ hypothesis, if we compute a value smaller than or equal to $1/\sqrt{3} + 0.5$ for $p = 1$, and smaller than or equal to $\frac{2}{3} + 0.5$ for $p > 1$, and to decide in favor of the $I(1)$ hypothesis if we get a larger value, gives an almost sure rule for choosing between $I(0)$ and $I(1)$.

NOTE

1. It should be mentioned that strong consistent selection procedures for choosing between competing models (e.g., lag order selection in ARMA models) have been already proposed (e.g., Hannan, 1980; Pötscher, 1983).

REFERENCES

Andrews, D.W.K. (1984) Non-strong mixing autoregressive processes. *Journal of Applied Probability* 21, 930–934.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Andrews, D.W.K. & J.C. Monahan (1992) An improved heteroskedasticity and autocorrelation consistent variance estimator. *Econometrica* 60, 953–966.

Burkholder, D.L. (1973) Distribution function inequalities for martingales. *Annals of Probability* 1, 19–42.

Davidson, J. (1994) *Stochastic Limit Theory.* New York: Oxford University Press.

Eberlain, E. (1986) On strong invariance principles under dependence assumptions. *Annals of Probability* 14, 260–270.

Gallant, A.R. & H. White (1988) *A Unified Theory for Estimation and Inference for Nonlinear Econometric Models.* New York: Basil Blackwell.

Hannan, E.J. (1980) The estimation of the order of an ARMA process. *Annals of Statistics* 8, 1071–1081.

Hansen, B.E. (1992) Consistent covariance matrix estimation for dependent heterogeneous processes. *Econometrica* 60, 967–972.

Karatzas, I. & S.E. Shreve (1991) *Brownian Motion and Stochastic Calculus*, 2nd ed. New York: Springer Verlag.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–178.

McKean, H.P. (1969) *Stochastic Integrals.* New York: Academic Press.

McLeish, D.L. (1975a) Invariance principles for dependent variables. *Probability Theory and Related Fields* 32, 165–178.

McLeish, D.L. (1975b) A maximal inequality and dependent strong laws. *Annals of Probability* 3, 829–839.

Newey, W.K. & K.D. West (1987) A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.

Newey, W.K. & K.D. West (1994) Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61, 631–653.

Phillips, P.C.B. & W. Ploberger (1994) Posterior odds testing for a unit root with data-based model selection. *Econometric Theory* 10, 774–808.

Phillips, P.C.B. & W. Ploberger (1996) An asymptotic theory of Bayesian inference for time series. *Econometrica* 64, 381–412.

Pötscher, B.M. (1983) Order estimation in ARMA models by Lagrange multiplier tests. *Annals of Statistics* 11, 872–885.

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

Robinson, P.M. (1991) Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* 1329–1363.
Serfling, R.J. (1968) Contributions to central limit theory for dependent variables. *Annals of Mathematical Statistics* 39, 1158–1175.
Sin, C.Y. & H. White (1996) Information criteria for selecting possibly mispecified parametric models. *Journal of Econometrics* 71, 207–225.
Sin, C.Y. (1996) On Selecting Strictly Nested Linear Regression Models with Information Criteria. Mimeo, City University of Hong Kong.
Stock, J.H. (1994) Deciding between $I(1)$ and $I(0)$. *Journal of Econometrics* 63, 105–131.
Strassen, V. (1964) An invariance principle for the law of the iterated logarithm. *Probability Theory and Related Fields* 3, 211–226.
White, H. (1984) *Asymptotic Theory for Econometricians.* New York: Academic Press.

APPENDIX

Hereafter let $\|u_t\|_r = (\sum_{i=1}^{p} E(|u_{it}|^r))^{1/r}$ and let $b_T = T^{3/2}\sqrt{2\log\log T}$ and $d_T = \sqrt{2T\log\log T}$.

Proof of Lemma 2.1.

(i) Given A1–A3, the result in (i) follows from Theorem 1 in Eberlain (1986), once we have shown (A.1) and (A.2), which follow.

$$\|E(S_T(m)|F_m)\|_2 \leq C \tag{A.1}$$

for C independent of m and $F_m = \sigma(\ldots\epsilon_{m-1},\epsilon_m)$. Here $\forall i,j = 1,2,\ldots p$, uniformly in m,

$$T^{-1}E|E(S_{iT}(m)S_{jT}(m)|F_m) - E(S_{iT}(m)S_{jT}(m))| = O(T^{-\theta}) \tag{A.2}$$

for some $\theta > 0$.

We begin by showing (A.1). By Lemma 3.5 in McLeish (1975a) and by Minkowski's inequality,

$$\begin{aligned}\|E(S_T(m)|F_m)\|_2 &\leq \|E(\epsilon_{m+1}|F_m)\|_2 + \|E(\epsilon_{m+2}|F_m)\|_2 + \cdots + \|E(\epsilon_{m+T}|F_m)\|_2 \\ &\leq 6\alpha_1^{1/2-1/2(4+\delta)}\|\epsilon_{m+1}\|_{2(4+\delta)} + \cdots \\ &\quad + 6\alpha_T^{1/2-1/2(4+\delta)}\|\epsilon_{m+T}\|_{2(4+\delta)} \\ &\leq 6\sup_t\|\epsilon_t\|_{2(4+\delta)}\sum_{j=0}^{\infty}\alpha_j^{1/2-1/2(4+\delta)} \leq 6pC_1^{1/2(4+\delta)}A \leq C\end{aligned}$$

because of the bound in A1 and because, given A2, $\sum_{j=1}^{\infty}\alpha_j^{1/2-1/2(4+\delta)} = A < \infty$.

It remains to prove (A.2). The next few steps are close to Serfling's (1968, pp. 1163–1164), but we allow for $i \neq j$, whereas Serfling considers the scalar case.

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

Recall that $S_{iT}(m) = \sum_{k=m+1}^{m+T} \epsilon_k$ and $S_{iT}(m+a) = \sum_{k=m+a+1}^{m+a+T} \epsilon_k$. The left hand side (LHS) of (A.2) can be rewritten as

$$\begin{aligned}
&E|E(T^{-1}S_{iT}(m+a)S_{jT}(m+a)|F_m) - E(T^{-1}S_{iT}(m+a)S_{jT}(m+a)) \\
&\quad - E((T^{-1}S_{iT}(m+a)S_{jT}(m+a) - T^{-1}S_{iT}(m)S_{jT}(m))|F_m) \\
&\quad + E(T^{-1}S_{iT}(m+a)S_{jT}(m+a) - T^{-1}S_{iT}(m)S_{jT}(m))| \\
&\leq E|E(T^{-1}S_{iT}(m+a)S_{jT}(m+a)|F_m) \\
&\quad - E(T^{-1}S_{iT}(m+a)S_{jT}(m+a))| \\
&\quad + 2E|T^{-1}S_{iT}(m+a)S_{jT}(m+a) - T^{-1}S_{iT}(m)S_{jT}(m)| \qquad \textbf{(A.3)}
\end{aligned}$$

We shall show that the first term on the right hand side (RHS) of the inequality in (A.3) is $O(T^{-\theta_1})$ and the second term is $O(T^{-\theta_2}), \theta_1, \theta_2 > 0$. Equation (A.2) then follows by setting $\theta = \text{Min}\{\theta_1, \theta_2\}$. As for the first term on the RHS of (A.3), from McLeish mixing inequality (1975a, Lemma 3.5), Cauchy–Schwartz and Minkowski inequality, we have

$$\begin{aligned}
&T^{-1}E|E(S_{iT}(m+a)S_{jT}(m+a)|F_m) - E(S_{iT}(m+a)S_{jT}(m+a))| \\
&\quad \leq 6T^{-1}\alpha_a^{1-(1/(2+\delta))}\|S_{iT}(m+a)S_{jT}(m+a)\|_{2+\delta} \\
&\quad \leq 6T^{-1}\alpha_a^{(1+\delta)/(2+\delta)}\|S_{iT}(m+a)\|_{2(2+\delta)}\|S_{jT}(m+a)\|_{2(2+\delta)} \\
&\quad \leq 6T^{-1}T^2\alpha_a^{(1+\delta)/(2+\delta)}(\sup_t E(\epsilon_{it}^{2(2+\delta)}))^{1/(2+\delta)} \\
&\quad \leq 6C_1^{1/(2+\delta)}T\alpha_a^{(1+\delta)/(2+\delta)}
\end{aligned}$$

where C_1 is the bound in A1. Now, given the size condition in A2, for $a = T^{\phi}$, with $[\delta(2+\delta)]/[4(\delta+1)(4+\delta)] < \phi < 1$, $T\alpha_a^{(\delta+1)/(2+\delta)}$ will approach zero, as $T \to \infty$ at rate $T^{-\theta_1}$, with $\theta_1 = \{\phi[4(4+\delta)(1+\delta)]/[\delta(2+\delta)]\} - 1$.

We now need to show that by setting $a = T^{\phi}$, with ϕ defined as previously, the second term on the RHS of (A.3) is $O(T^{-\theta_2}), \theta_2 > 0$. First note that

$$\begin{aligned}
&T^{-1}S_{iT}(m+a)S_{jT}(m+a) - T^{-1}S_{iT}(m)S_{jT}(m) \\
&\quad = (T^{-1/2}S_{jT}(m+a) - T^{-1/2}S_{jT}(m))(T^{-1/2}S_{iT}(m+a) + T^{-1/2}S_{iT}(m)) \\
&\quad\quad + T^{-1/2}S_{jT}(m+a)(T^{-1/2}S_{iT}(m+a) - T^{-1/2}S_{iT}(m)) \\
&\quad\quad - T^{-1/2}S_{iT}(m+a)(T^{-1/2}S_{jT}(m+a) - T^{-1/2}S_{jT}(m)).
\end{aligned}$$

Needless to say, the sum of the last two terms is identical to zero, for $i = j$. Thus, by the triangular inequality and Cauchy–Schwartz inequality,

$$\begin{aligned}
&E|T^{-1}S_{jT}(m+a)S_{iT}(m+a) - T^{-1}S_{jT}(m)S_{iT}(m)| \\
&\quad \leq (E(T^{-1/2}S_{jT}(m+a) - T^{-1/2}S_{jT}(m))^2)^{1/2} \\
&\quad\quad \times (E(T^{-1/2}S_{iT}(m+a) + T^{-1/2}E(S_{jT}(m))^2)^{1/2} \\
&\quad\quad + (E(T^{-1/2}S_{jT}(m+a))^2)^{1/2}(E(T^{-1/2}S_{iT}(m+a) - T^{-1/2}S_{iT}(m))^2)^{1/2} \\
&\quad\quad + (E(T^{-1/2}S_{iT}(m+a))^2)^{1/2}(E(T^{-1/2}S_{jT}(m+a) - T^{-1/2}S_{jT}(m))^2)^{1/2} \\
&\quad = O(a^{1/2}T^{-1/2}),
\end{aligned}$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

as, given A3, $E(T^{-1/2}S_{iT}(m+a) - T^{-1/2}S_{iT}(m))^2 = O(aT^{-1})$, and $E(T^{-1/2}S_{iT}(m+a))^2 = O(1)$, uniformly in m, for $a = o(T)$. Thus by setting $a = T^{\phi}$, we have that the second term in (A.3) is $O(T^{-\theta_2})$, with $\theta_2 = \frac{1}{2}(1-\phi)$. This completes the proof of (i).

(ii) Given the statement in (i) we have that

$$d_T^{-1}\Omega_0^{-1/2}\sum_{j=1}^{T}\epsilon_j - d_T^{-1}W(T) = o_{as}(1).$$

Recalling that ϵ_t is R^p-valued and p is finite, we also have that

$$d_T^{-1}\left\|\Omega_0^{-1/2}\sum_{j=1}^{T}\epsilon_j - W(T)\right\| = o_{as}(1),$$

and so

$$d_T^{-1}\left\|\Omega_0^{-1/2}\sum_{j=1}^{T}\epsilon_j\right\| = d_T^{-1}\|W(T)\| + o_{as}(1). \tag{A.4}$$

From the law of the iterated logarithm (LIL) for the p-dimensional Brownian motion (e.g., McKean, 1969, p. 96), we know that $\limsup_{T\to\infty} d_T^{-1}\|W(T)\| = 1$, with probability one. The statement in (ii) then follows from (A.4). ■

Proof of Corollary 2.2. From the proof of Lemma 2.1(i), we have seen that

$$\|E(S_T(m)|F_m)\|_2 \le C.$$

It then follows from the first part of Theorem 2 in Eberlain (1986) that A3 is satisfied. ■

Proof of Lemma 3.1. We first consider the case of $p = 1$, as the generalization to arbitrary finite p is straightforward.

We begin by showing that

$$E\left(\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T}(\epsilon_t\epsilon_{t-\tau} - E(\epsilon_t\epsilon_{t-\tau}))\right)^4 = O(T^{-(1+\eta)}) \tag{A.5}$$

for some $\eta > 0$. We then show that (A.5) implies the desired result. For notational convenience let

$$z_{t,\tau} = \epsilon_t\epsilon_{t-\tau} - E(\epsilon_t\epsilon_{t-\tau})$$

and

$$Z_{T,\tau} = \sum_{t=\tau+1}^{T} z_{t,\tau}.$$

From Loeve's c_r inequality (e.g., Davidson, 1994, p. 140),

$$E\left(\sum_{\tau=1}^{l_T} w_\tau Z_{T,\tau}\right)^4 \le l_T^3\sum_{\tau=1}^{l_T}(E|\omega_\tau Z_{T,\tau}|^4) \le l_T^4 \sup_\tau E|Z_{T,\tau}|^4. \tag{A.6}$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

Now we shall found a bound for $E|Z_{T,\tau}|^4$ for a generic τ, and then we shall take the supremum over $\tau = 1,2,\ldots,l_T$. Hereafter let

$$E_{t-k}z_{t,\tau} = E(z_{t,\tau}|G_{t-k}),$$

where $G_{t-k} = \sigma(\ldots z_{t-k-1,\tau}, z_{t-k,\tau})$. We have

$$z_{t,\tau} = \sum_{k=-\infty}^{\infty} (E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}) = \sum_{k=0}^{\infty} (E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}).$$

In fact, for $k < 0$, $E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau} = z_{t,\tau} - z_{t,\tau} = 0$, and for $k \to \infty$, $E_{t-k}z_{t,\tau}$ converges almost surely to $E(z_{t,\tau}) = 0$.

We can write

$$Z_{T,\tau} = \sum_{t=\tau+1}^{T} z_{t,\tau} = \sum_{t=\tau+1}^{T} \sum_{k=0}^{\infty} (E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}) = \sum_{k=0}^{\infty} Y_{Tk,\tau}, \tag{A.7}$$

where

$$Y_{Tk,\tau} = \sum_{t=\tau+1}^{T} (E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}).$$

Now let $K_\tau = \sum_{k=0}^{\infty} \gamma_k/\tau$, where $\gamma_k > 0$, $\gamma_k = 1$ for $k = 0,1,\ldots\tau$, and $\gamma_k = o(k^{-1})$ for $k > \tau$; thus $\sup_{\tau=1,\ldots l_T} K_\tau = L < \infty$. Following Lemma 6.1 and Lemma 6.2 in McLeish (1975a), we have that

$$E|Z_{T,\tau}|^4 = E\left|\sum_{k=0}^{\infty} Y_{Tk,\tau}\right|^4 = K_\tau^4 E\left|\sum_{k=0}^{\infty} \frac{Y_{Tk,\tau}}{\gamma_k/\tau}\frac{\gamma_k/\tau}{K_\tau}\right|^4$$

$$\leq K_\tau^3 \sum_{k=0}^{\infty} \left(\frac{\gamma_k}{\tau}\right)^{-3} E|Y_{Tk,\tau}|^4. \tag{A.8}$$

The next few steps are close to those followed by Davidson (1994) in the proof of his Theorem 16.11. However he treats the case of the rth moment, $r < 2$, whereas we consider the fourth moment, and we also need to take care of the fact that for $0 \leq k \leq \tau$, $z_{t,\tau}$, and $z_{t-k,\tau}$ belong to overlapping σ-fields.

From Burkholder's inequality and from Loeve's c_r inequality, we have that

$$E|Y_{Tk,\tau}|^4 \leq C_4^4 E\left|\sum_{t=\tau+1}^{T} (E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau})^2\right|^2$$

$$\leq C_4^4 T \sum_{t=\tau+1}^{T} E|E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}|^4, \tag{A.9}$$

where $C_4 \leq 18rq^{1/2}$, $r = 4$, and $q^{-1} = 1 - r^{-1}$ (see Burkholder, 1973, p. 23).

To apply mixing inequalities, we need to define the mixing coefficients of $z_{t,\tau}$. Recall that $G_a^b = \sigma(z_{a,\tau},\ldots,z_{b,\tau})$, $z_{a,\tau} = \epsilon_a\epsilon_{a-\tau} - E(\epsilon_a\epsilon_{a-\tau}) = g(\epsilon_a,\epsilon_{a-\tau})$, $F_a^b = \sigma(\epsilon_a\ldots\epsilon_b)$. Following the lines of White (1984, Corollary 6.18), $\alpha_k^\tau = \sup_n \alpha(G_{-\infty}^n, G_{n+k}^\infty)$, $\alpha_k = \sup_n \alpha(F_{-\infty}^n, F_{n+k}^\infty)$, and so $\alpha_{k-\tau} = \sup_n \alpha(F_{-\infty}^n, F_{n+k-\tau}^\infty)$. As g is measurable, $G_a^b \subset F_{a-\tau}^b$,

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

so $\alpha_k^\tau \le \sup_n \alpha(F_\infty^n, F_{n+k-\tau}^\infty) = \alpha_{k-\tau}$. It then follows that $\alpha_k^\tau \le 1$, $0 \le k \le \tau$, and $\alpha_k^\tau \le \alpha_{k-\tau}$, $k > \tau$.

From Minkowski's inequality and from McLeish mixing inequality (McLeish, 1975a, Lemma 3.5), for $k > \tau$ we have

$$\begin{aligned}\|E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}\|_4 &\le \|E_{t-k}z_{t,\tau}\|_4 + \|E_{t-k-1}z_{t,\tau}\|_4 \\ &\le 2(2^{1/4}+1)(\alpha_k^\tau)^{1/4-1/(4+\delta)}\|z_{t,\tau}\|_{4+\delta} \\ &\le 2(2^{1/4}+1)\alpha_{k-\tau}^{\delta/(4(4+\delta))}\|z_{t,\tau}\|_{4+\delta}. \end{aligned} \tag{A.10}$$

Now by Minkowski and Cauchy–Schwartz inequality, we have

$$\begin{aligned}\|z_{t,\tau}\|_{4+\delta} &= \|\epsilon_t\epsilon_{t-\tau} - E(\epsilon_t\epsilon_{t-\tau})\|_{4+\delta} \le \|\epsilon_t\epsilon_{t-\tau}\|_{4+\delta} + E(|\epsilon_t\epsilon_{t-\tau}|) \\ &\le \sup_t (E|\epsilon_t|^{2(4+\delta)})^{1/(2(4+\delta))} \sup_t (E|\epsilon_t|^{2(4+\delta)})^{1/(2(4+\delta))} \\ &\quad + \sup_t (E\epsilon_t^2)^{1/2} \sup_t (E\epsilon_t^2)^{1/2} \le 2C_1^{1/(4+\delta)},\end{aligned}$$

where C_1 is the bound in A1. Thus for $k > \tau$,

$$\|E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}\|_4 \le 4(2^{1/4}+1)C_1^{1/(4+\delta)}\alpha_{k-\tau}^{\delta/(4(4+\delta))}.$$

Now consider the case of $0 \le k \le \tau$, in which $\alpha_k^\tau \le 1$:

$$\|E_{t-k}z_{t,\tau} - E_{t-k-1}z_{t,\tau}\|_4 \le 2\|E_{t-k-1}z_{t,\tau}\|_4 \le 4(2^{1/4}+1)C_1^{1/(4+\delta)}.$$

Thus,

$$E(Y_{Tk,\tau})^4 \le CT^2 \quad \text{for } 0 \le k \le \tau \qquad E(Y_{Tk,\tau})^4 \le CT^2(\alpha_{k-\tau}^{\delta/(4(4+\delta))})^4 \quad \text{for } k > \tau, \tag{A.11}$$

where $C = 16(2^{1/4}+1)^4 C_4^4 C_1^{4/(4+\delta)}$, with C_4 defined as following (A.9).

From (A.8) and (A.11) we can write

$$E|Z_{T,\tau}|^4 \le CT^2\left(\sum_{k=0}^{\infty}\frac{\gamma_k}{\tau}\right)^3\tau^4\left(\sum_{k=0}^{\tau}\left(\frac{\gamma_k}{\tau}\right)^{-3}\frac{1}{\tau_4} + \sum_{k=\tau+1}^{\infty}\left(\frac{\gamma_k}{\tau}\right)^{-3}\left(\frac{\alpha_k^{\delta/(4(4+\delta))}}{\tau}\right)^4\right). \tag{A.12}$$

We can set $\gamma_k = 1$, $k = 0, \ldots \tau$, and $\gamma_k = \alpha_{k-\tau}^{\delta/(4(4+\delta))}$, $k > \tau$, because given A2, $\alpha_{k-\tau}^{\delta/(4(4+\delta))} = o((k-\tau)^{-1}) = o(k^{-1})$. Thus

$$E|Z_{T,\tau}|^4 \le CL^3T^2\tau^4 + CL^3T^2\tau^4\left(\sum_{k=\tau+1}^{\infty}\frac{\alpha_{k-\tau}^{\delta/(4(4+\delta))}}{\tau}\right), \tag{A.13}$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

where $L = \sup_{\tau=1,\dots l_T} \sum_{k=0}^{\infty}(\gamma_k/\tau)$. Given the size condition in A2, we have that

$$\sup_{\tau=1,2,\dots l_T} \sum_{k=\tau+1}^{\infty} \frac{\alpha_k^{\delta/(4(4+\delta))}}{\tau} \leq B < \infty$$

for some constant B. Thus

$$\sup_{\tau=1,2,\dots l_T} E|Z_{T,\tau}|^4 \leq T^2 l_T^4 C L^3 (1+B).$$

Thus, from (A.6), we have that

$$E\left(\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} z_{t,\tau}\right)^4 = O(T^2 l_T^8),$$

and for $l_T = o(T^{1/8})$, we have

$$E\left(\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} z_{t,\tau}\right)^4 = O(T^{-(1+\eta)})$$

for $\eta > 0$. This proves (A.5).

Now we need to show that (A.5) implies the desired result. Let

$$A_T = \left\{\omega : \left| T^{-1} \sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} z_{t,\tau}(\omega) \right| > \epsilon \right\}$$

for some $\epsilon > 0$. By Markov inequality we have that

$$P(A_T) \leq \frac{E\left(T^{-1}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} z_{t,\tau}\right)^4}{\epsilon^4} = O(T^{-(1+\eta)}).$$

Thus $\sum_{T=1}^{\infty} P(A_T) < \infty$. From the first Borel Cantelli lemma we then have that

$$P\left(\limsup_T A_T\right) = 0.$$

This concludes the proof for the case of $p = 1$.

For the case of arbitrary finite p, as almost sure convergence of all the elements implies almost sure convergence of the matrix, it suffices to show that for all $i, j = 1,2,\dots p$,

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=1}^{T-\tau} (\epsilon_{it}\epsilon_{jt-\tau} - E(\epsilon_{it}\epsilon_{jt-\tau})) = o_{as}(1).$$

This can be shown by the same argument used previously. ■

Proof of Theorem 3.2. Hereafter $\ddot{\Omega}_T$, $\tilde{\Omega}_T$, $\bar{\Omega}_T$, Ω_0, and $\hat{\Omega}_T$ are defined as at the beginning of Section 3. We first show that $\ddot{\Omega}_T - \Omega_0 = o_{as}(1)$. By Lemma 6.17 in White (1984), if, as $T \to \infty$, $l_T \to \infty$, then $\tilde{\Omega}_T - \bar{\Omega}_T \to 0$. As, given A3, $\bar{\Omega}_T \to \Omega_0$, it suffices to show that $\ddot{\Omega}_T - \tilde{\Omega}_T = o_{as}(1)$. As almost sure convergence element by element implies almost sure

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

convergence of the matrix, it suffices to show that $\ddot{\omega}_{ij,T} - \tilde{\omega}_{ij,T} = o_{as}(1)$, $\forall i, j = 1,2,\ldots p$. We can write

$$\begin{aligned}\ddot{\omega}_{ij,T} - \tilde{\omega}_{ij,T} = \frac{1}{T}\sum_{t=1}^{T}(\epsilon_{it}\epsilon_{jt} - E(\epsilon_{it}\epsilon_{jt}))\\ + \frac{1}{T}\sum_{\tau}^{T} w_\tau \sum_{t=\tau+1}^{T}((\epsilon_{it}\epsilon_{jt-\tau} - E(\epsilon_{ti}\epsilon_{jt-\tau})) + (\epsilon_{jt}\epsilon_{it-\tau} - E(\epsilon_{jt}\epsilon_{it-\tau})))\\ + \frac{1}{T}\sum_{\tau=1}^{l_T}(w_\tau - 1)\left(\sum_{t=\tau+1}^{T}(E(\epsilon_{it}\epsilon_{jt-\tau}) + E(\epsilon_{jt}\epsilon_{it-\tau}))\right). \qquad \textbf{(A.14)}\end{aligned}$$

The last term converges to zero by the same argument used in the proof of Theorem 2 in Newey and West (1987). Given A1 and A2, the first term on the RHS of (A.14) is $o_{as}(1)$ by the strong law of large numbers (e.g., McLeish, 1975b, Theorem 2.10). Finally the second term is $o_{as}(1)$ because of Lemma 3.1.

Now it remains to show that $\hat{\Omega}_T - \ddot{\Omega}_T = o_{as}(1)$,

$$\begin{aligned}\hat{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T}((X_t - \mu) - (\bar{X} - \mu))((X_t - \mu) - (\bar{X} - \mu))'\\ + \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T}[((X_t - \mu) - (\bar{X} - \mu))((X_{t-\tau} - \mu) - (\bar{X} - \mu))'\\ + ((X_{t-\tau} - \mu) - (\bar{X} - \mu))((X_t - \mu) - (\bar{X} - \mu))']\\ = \ddot{\Omega}_T + (\bar{X} - \mu)(\bar{X} - \mu)' - (\bar{X} - \mu)\frac{1}{T}\sum_{t=1}^{T}(X_t - \mu)' + \frac{1}{T}\sum_{t=1}^{T}(X_t - \mu)(\bar{X} - \mu)'\\ - (\bar{X} - \mu)\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T}((X_{t-\tau} - \mu)' + (X_t - \mu)')\\ - \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T}((X_t - \mu) + (X_{t-\tau} - \mu))(\bar{X} - \mu)'\\ + 2(\bar{X} - \mu)(\bar{X} - \mu)'\sum_{\tau}^{l_T} w_\tau \frac{T - \tau}{T} = \ddot{\Omega}_T + I + II + III + IV + V + VI.\end{aligned}$$

We want to show that $I + II + III + IV + V + VI = o_{as}(1)$. The terms I, II, III are $o_{as}(1)$ because of the strong law of large numbers for strong mixing. Now $(\bar{X} - \mu) = 1/T\sum_{t=1}^{T}(X_t - \mu) = O_{as}(\sqrt{2\log\log T}/\sqrt{T})$ by Lemma 2.1(ii), and as $l_T = o(T^{1/8})$ it follows that VI is $o_{as}(1)$. Finally given the conditions on the rate of growth of l_T and given A1–A3, IV and V are $o_{as}(1)$ by Lemma 2.1(ii) and by Lemma 3.1. ■

Proof of Corollary 3.3. It follows by the same argument used in the proof of Lemma 3.1, by noting that (A.5) should be modified as

$$E\left(T^{-1}\sum_{\tau=1}^{l_T} w_\tau Z_{t,\tau}\right)^{2\beta} = O(T^{-(1+\eta)})$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

with $\eta > 0$. Equation (A.6) should be modified as

$$E\left(\sum_{\tau=1}^{l_T} w_\tau Z_{t,\tau}\right)^{2\beta} \leq l_T^{2\beta} \sup_\tau E|Z_{T,\tau}|^{2\beta},$$

(A.9) should be modified as

$$E|Y_{Tk,\tau}B|^{2\beta} \leq C_{2\beta}^{2\beta} T^{\beta-1} \sum_{t=1}^{T-\tau} E|E_{t-k} z_{t,\tau} - E_{t-k-1} z_{t,\tau}|^{2\beta},$$

and so we have

$$\sup_\tau E|Z_{T,\tau}|^{2\beta} \leq CT^\beta l_T^{2\beta}$$

for some $0 < C < \infty$, and it follows that we need $l_T = o(T^{\beta - 1/4\beta})$. ■

Proof of Theorem 4.1. Most of this proof is the same for $p = 1$ and $p > 1$. We shall just stress the specific points in which the one-dimensional and the multidimensional case require a different treatment. Without loss of generality set $r_0 = 0$, where r_t is defined in (2).

$$\begin{aligned}
&\frac{1}{b_T}\sum_{t=1}^{T}\left\|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j - \bar{X})\right\| \\
&\quad = \frac{1}{b_T}\sum_{t=1}^{T}\left\|\Omega_0^{-1/2}\sum_{j=1}^{t}(X_j-\mu) - \Omega_0^{-1/2}t(\bar{X}-\mu) + (\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2})\right. \\
&\qquad\qquad \left.\times \sum_{j=1}^{t}(X_j-\mu) - (\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2})t(\bar{X}-\mu)\right\| \\
&\quad \leq \frac{1}{b_T}\sum_{t=1}^{T}\left\|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\right\| + \frac{1}{Tb_T}\sum_{t=1}^{T} t\left\|\Omega_0^{-1/2}\sum_{s=1}^{T}\epsilon_s\right\| \\
&\qquad + \frac{1}{b_T}\sum_{t=1}^{T}\left\|(\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2})\sum_{j=1}^{t}\epsilon_j\right\| + \frac{1}{Tb_T}\sum_{t=1}^{T} t\left\|(\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2})\sum_{s=1}^{T}\epsilon_s\right\|. \quad \textbf{(A.15)}
\end{aligned}$$

We shall organize the proof in three different steps.

Step 1. We show that, for $p = 1$, almost surely

$$\text{(a)}\ \limsup_{T\to\infty} \frac{1}{T^{3/2}\sqrt{2\log\log(T)}}\sum_{t=1}^{T}\left\|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\right\| = \frac{1}{\sqrt{3}}$$

and for $p > 1$, almost surely,

$$\text{(b)}\ \limsup_{T\to\infty} \frac{1}{T^{3/2}\sqrt{2\log\log(T)}}\sum_{t=1}^{T}\left\|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\right\| \leq \frac{2}{3}.$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

Let $t = [\mathrm{Tr}]$ and let $\sum_{j=1}^{[\mathrm{Tr}]} \Omega_0^{-1/2} \epsilon_j = S_{[\mathrm{Tr}]} = S_T(r)$, and also let

$$\eta_T(r) = ([\mathrm{Tr}] + 1 - \mathrm{Tr}) S_T(r) + (\mathrm{Tr} - [\mathrm{Tr}]) S_{T+1}(r).$$

From Lemma 2.1(i), for any finite p, as $T \to \infty$,

$$\sup_{r \in [0,1]} \frac{1}{d_T} \| S_{[\mathrm{Tr}]} - W(\mathrm{Tr}) \| \xrightarrow{\text{a.s.}} 0,$$

and, as

$$\sup_{r \in [0,1]} \frac{1}{d_T} \| \eta_{(\mathrm{Tr})} - S_{[\mathrm{Tr}]} \| \xrightarrow{\text{a.s.}} 0,$$

it follows that the set of norm limit points of $\{d_T^{-1} S_T\}$ and $\{d_T^{-1} \eta_T\}$ coincides with the set of norm limit points of $d_T^{-1} W_T$ with probability one. From Theorem 1 in Strassen (1964), we know that $\{d_T^{-1} W_T, T \geq 3\}$ is relatively norm compact and the set of its limit points coincides, almost surely, with K, where $K = \{ f \in [0,1] \to R^p;\ f(0) = 0, f \text{ absolutely continuous}, \int_0^1 \| df(r)/dr \|^2 \, dr \leq 1 \}$.

We now make use of the following result, due to Strassen (1964, p. 218).

LEMMA A. *If ϕ is a continuous map from C, the space of continuous functions on $[0,1]$, closed with respect to the sup norm, to some Hausdorf space, e.g., a Euclidean space as in our case, then, with probability one, $\{\phi(d_T^{-1} \eta_T), T \geq 3\}$, where η_T is defined as previously, is relatively norm compact and the set of norm limit points coincides, almost surely, with $\phi(K)$, with K defined as previously.* ■

Now,

$$P\left[\limsup_{T \to \infty} \frac{1}{b_T} \sum_{t=1}^{T} \left\| \Omega_0^{-1/2} \sum_{j=1}^{t} \epsilon_j \right\| = \limsup_{T \to \infty} d_T^{-1} \int_0^1 \| \eta_T(t) \| \, dt \right] = 1.$$

From Lemma A, for $p = 1$, we have

$$P\left[\limsup_{T \to \infty} d_T^{-1} \int_0^1 \| \eta_T(t) \| \, dt = \sup \phi(K) \right] = 1.$$

Given the definition of K, following Strassen (1964, pp. 218–219),

$$\sup \phi(K) = \sup_{u \in K} \int_0^1 \| u_t \| \, dt = \sup_{u \in K} \int_0^1 \| \dot{u}_t \| (1 - t) \, dt$$

$$= \sup \left[\int_0^1 (1 - t) \| y(t) \| \, dt : \int_0^1 \| y(t) \|^2 \, dt \leq 1 \right] = \left(\int_0^1 (1 - t)^2 \, dt \right)^{1/2},$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

where the second equality on the RHS comes from integration by parts and the last equality follows from the evaluation of the supremum of a linear functional on the unit sphere of a Hilbert space. As

$$\left(\int_0^1 (1-t)^2\,dt\right)^{1/2} = \frac{1}{\sqrt{3}},$$

the statement of step 1(a) follows.

The same argument does not apply when $p > 1$, as $d\|u_t\|/dt \neq \|du_t/dt\|$. Now for arbitrary finite p,

$$\sup \phi(K) = \sup_{u\in K} \int_0^1 \|u_t\|\,dt.$$

Now

$$\sup_{u\in K} \int_0^1 \|u_t\|\,dt = \sup_{u\in K} \int_0^1 \left\| \int_0^t \dot{u}_s\,ds \right\| dt \leq \int_0^1 \left(\sup_{u\in K} \left(\int_0^t ds \int_0^t \dot{u}_{s'}\dot{u}_s\,ds \right)^{1/2} \right) dt. \quad \textbf{(A.16)}$$

By the definition of K,

$$\int_0^1 \dot{u}_{s'}\dot{u}_s\,ds \leq 1.$$

As the bound does not depend on the specific u we pick, we also have that

$$\sup_{u\in K} \int_0^1 \dot{u}_{s'}\dot{u}_s\,ds \leq 1.$$

Thus the RHS in (A.16) is majorized by $\int_0^1 t^{1/2}\,dt = \frac{2}{3}$. The statement in step 1(b) then follows.

Step 2. For arbitrary finite p,

$$\limsup_{T\to\infty} \frac{1}{b_T} \sum_{t=1}^{T} t\|\Omega_0^{-1/2}(\bar{X}-\mu)\|$$

$$= \limsup_{T\to\infty} \frac{T(T+1)}{2T^2} \frac{1}{\sqrt{2T\log\log(T)}} \left\| \Omega_0^{-1/2} \sum_{t=1}^{T} \epsilon_t \right\| = \frac{1}{2},$$

recalling that by Lemma 2.1(ii)

$$\limsup_{T\to\infty} \frac{1}{\sqrt{2T\log\log(T)}} \left\| \Omega_0^{-1/2} \sum_{t=1}^{T} \epsilon_t \right\| = 1.$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

Step 3. It remains to show that the last two terms on the RHS of (A.15) are $o_{as}(1)$. Let $\hat{\Omega}_T^{-1/2} = [\hat{u}_{ij}]$ and $\Omega_0^{-1/2} = [u_{ij}], \forall i, j = 1,2,\ldots p$.

$$\frac{1}{b_T}\sum_{t=1}^{T}\left\|(\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2})\sum_{j=1}^{t}\epsilon_j\right\|$$

$$\leq \frac{1}{b_T}\sum_{t=1}^{T}\left|\sum_{i=1}^{p}(\hat{u}_{i1} - u_{i1})\sum_{j=1}^{t}\epsilon_{1j}\right| + \cdots + \frac{1}{b_T}\sum_{t=1}^{T}\left|\sum_{i=1}^{p}(\hat{u}_{ip} - u_{ip})\sum_{j=1}^{t}\epsilon_{pj}\right| = o_{as}(1),$$

because of Theorem 3.2 and because, by step 1, $b_T^{-1}\sum_{t=1}^{T}\sum_{j=1}^{t}\epsilon_{ij} = O_{as}(1)$, $\forall i = 1,\ldots p$. As for the last term on the RHS of (A.15), it can be majorized by

$$\frac{1}{T^2}\frac{T(T+1)}{2}\frac{1}{\sqrt{2T\log\log(T)}}\left|\sum_{i=1}^{p}(\hat{u}_{i1} - u_{i1})\sum_{t=1}^{T}\epsilon_{1t}\right|$$

$$+ \cdots + \frac{T(T+1)}{2T^2}\frac{1}{\sqrt{2T\log\log(T)}}\left|\sum_{i=1}^{p}(\hat{u}_{ip} - u_{ip})\sum_{t=1}^{T}\epsilon_{pt}\right| = o_{as}(1)$$

as by Theorem 3.2 $\hat{u}_{ij} - u_{ij} = o_{as}(1)$, $\forall ij$ and because $\forall j = 1,2\ldots p$,

$$\limsup_{T\to\infty}\frac{1}{\sqrt{2T\log\log(T)}}\left|\sum_{t=1}^{T}\epsilon_{jt}\right| = O_{as}(1).$$

The desired result then follows by putting together steps 1–3 and by recalling that the lim sup of the sum is less than or equal to the sum of the lim sup's. ■

Proof of Theorem 4.2. We shall begin by showing that

$$\hat{\Omega}_T = O_{as}(2Tl_T\log\log T). \tag{A.17}$$

This implies that no element of $\hat{\Omega}_T^{-1/2}$, with probability one, can converge to zero at a rate faster than $(2Tl_T\log\log T)^{-1/2}$. We then show that

$$P\left(\lim_{T\to\infty}\frac{1}{T^{5/2}}\sum_{t=1}^{T}\left\|\sum_{j=1}^{t}(X_j - \bar{X})\right\| > 0\right) = 1. \tag{A.18}$$

The desired result then follows from (A.17) and (A.18).

Without loss of generality, we can assume that X_{1t} is an integrated process. Let $\hat{\omega}_{11,T}$ be the {11} element of $\hat{\Omega}_T$. We shall show that

$$\frac{1}{2Tl_T\log\log T}\hat{\omega}_{11,T} = O_{as}(1).$$

Again without loss of generality we can assume that in (2) $r_0 = 0$; so we can write

$$X_{1t} = \mu + \epsilon_{1t} + \sum_{j=1}^{t}\xi_{1j}$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

and

$$\bar{X}_1 = \mu + \frac{1}{T}\sum_{t=1}^{T}\epsilon_{1t} + \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t}\xi_{1j}.$$

Recall that ξ_t is i.i.d., by noting that because of Lemma 2.1(i) (for ϵ_{1t}) and because of Theorem 2 in Strassen (1964, for ξ_{1t}), for any $\phi > 0$, for $t \leq T$,

$$\frac{1}{T^{\phi}\sqrt{2T\log\log T}}\sum_{j=1}^{t}\epsilon_{1j} = o_{as}(1)$$

and

$$\frac{1}{T^{\phi}\sqrt{2T\log\log T}}\sum_{j=1}^{t}\xi_{1j} = o_{as}(1),$$

we have that

$$\begin{aligned}
&\frac{1}{Tl_T(2\log\log T)}\hat{\omega}_{11,T} \\
&\quad = \frac{2}{T^2 l_T(2\log\log T)}\sum_{\tau=1}^{l_T} w_\tau \\
&\qquad \times \sum_{t=\tau+1}^{T}\left(\sum_{j=1}^{t}\xi_{1j} - \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t}\xi_{1j}\right)\left(\sum_{j=1}^{t-\tau}\xi_{1j} - \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t}\xi_{1j}\right) + o_{as}(1) \\
&\quad = \frac{2}{T^2 l_T(2\log\log T)}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T}\left(\sum_{j=\tau}^{t}\xi_{1j} - \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t}\epsilon_{1j}\right)^2 + o_{as}(1) = I + o_{as}(1).
\end{aligned}$$

Now almost surely

$$\begin{aligned}
I &\leq \frac{2}{T^2(2\log\log T)}\sum_{t=\tau+1}^{T}\left(\sum_{j=1}^{t}\xi_{1j}\right)^2 + \left(\frac{6}{T^{3/2}\sqrt{2\log\log T}}\sum_{t=1}^{T}\left|\sum_{j=1}^{t}\xi_{1j}\right|\right)^2 \\
&\leq 2\,\frac{4\sigma_{11}^2}{\pi^2} + \left(6\,\frac{\sigma_{11}}{\sqrt{3}}\right)^2,
\end{aligned}$$

where the second inequality comes from Strassen (1964, p. 219) and σ_{11}^2 is the variance of ξ_1.

Suppose $X_{it}, i = 1\ldots r$, $1 \leq r \leq p$ is $I(1)$ and $X_{it}, i = r+1,\ldots p$ is $I(0)$; by the same argument given previously,

$$\frac{1}{Tl_T(2\log\log T)}\hat{\omega}_{ii,T} = O_{as}(1), \qquad i = 1,\ldots r$$

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press

and

$$\frac{1}{Tl_T(2\log\log T)}\,\hat{\omega}_{ij,T} = o_{as}(1), \forall i, \quad \text{and/or} \quad j = r+1\ldots p.$$

It follows that $\hat{\Omega}_T^{1/2} = O_{as}(\sqrt{2Tl_T \log\log T})$, and so with probability one no component of $\hat{\Omega}_T^{-1/2}$ can approach zero at a rate faster than $(2Tl_T \log\log T)^{-1/2}$. Now let $X_t = (X_t^{(r)}, X_t^{(p-r)})'$, where $X_t^{(r)}$ denotes the r-vector with $I(1)$ components and $X_t^{(p-r)}$ the $p-r$-vector of $I(0)$ components. We have

$$\begin{aligned}
&\frac{1}{T^{5/2}}\sum_{t=1}^{T}\left\|\sum_{j=1}^{t}(X_j - \bar{X})\right\| \\
&\quad = \frac{1}{T^{5/2}}\sum_{t=1}^{T}\left\|\sum_{j=1}^{t}(X_j^{(r)} - \bar{X}^{(r)})\right\| + o_{as}(1) \\
&\quad \geq \frac{1}{T^{5/2}}\sum_{t=1}^{T}\left|\sum_{j=1}^{t}(X_{1j} - \bar{X}_1)\right| + o_{as}(1) \\
&\quad = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{T}\sum_{j=1}^{t}\left(\frac{\sigma_{11}}{\sqrt{T}}\sum_{h=1}^{j}\frac{\xi_{1h}}{\sigma_{11}} - \frac{\sigma_{11}}{T^{3/2}}\sum_{t=1}^{T}\sum_{j=1}^{t}\frac{\xi_{1j}}{\sigma_{11}}\right)\right| \\
&\qquad + o_{as}(1).
\end{aligned}$$

From Theorem 2 in Eberlain (1986), we know that

$$\frac{1}{\sqrt{t}}\sum_{j=1}^{t}\frac{\xi_{1j}}{\sigma_{11}} - \frac{Wt}{\sqrt{t}} = O_{as}(t^{-\theta}), \quad \theta > 0,$$

where $W(t)$ is a standard one-dimensional Brownian motion.

Let us define the set of zeros of the Brownian motion as

$$Z = \{(t,\omega) \in [0,\infty) \times \Omega; W(t,\omega) = 0\}$$

and for a given $\omega \in \Omega$ define

$$Z(\omega) = \{t \in [0,\infty); W(t,\omega) = 0\}.$$

From Theorem 2.9.6 in Karatzas and Shreve (1991), $\forall \omega \in \Omega^*$, with $P(\Omega^*) = 1$, $Z(\omega)$ has Lebesgue measure zero. It then follows that

$$P\left(\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{T}\sum_{j=1}^{t}\left(\frac{1}{\sqrt{T}}\sum_{h=1}^{j}\xi_{1h} - \frac{1}{T^{3/2}}\sum_{t=1}^{T}\sum_{j=1}^{t}\xi_{1j}\right)\right| > 0\right) = 1,$$

and so

$$P\left(\lim_{T\to\infty}\frac{1}{T^{5/2}}\sum_{t=1}^{T}\left\|\sum_{j=1}^{t}(X_j - \bar{X})\right\| > 0\right) = 1.$$

Given (A.18), this concludes the proof. ■

https://doi.org/10.1017/S0266466699155014 Published online by Cambridge University Press