EXPLOITING CROSS SECTION VARIATION FOR UNIT ROOT INFERENCE IN DYNAMIC DATA

by

Danny Quah

DISCUSSION PAPER NO. 171

LSE FINANCIAL MARKETS GROUP

LONDON SCHOOL OF ECONOMICS

EXPLOITING CROSS SECTION VARIATION FOR UNIT ROOT INFERENCE IN DYNAMIC DATA

by

Danny Quah

DISCUSSION PAPER NO. 171

LSE FINANCIAL MARKETS GROUP

DISCUSSION PAPER SERIES

October 1993

Danny Quah is a Reader in Economics at the London School of Economics and a member of the Financial Markets Group. Any opinions expressed here are those of the author and not necessarily those of the Financial Markets Group.

Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data

by

Danny Quah
LSE Economics Department, FMG, and CEPR
Houghton Street
London WC2A 2AE
dquah@lse.ac.uk

October 1993

Abstract

This paper considers unit root regressions in data having simultaneously extensive cross-section and time-series variation. The standard least squares estimators in such data structures turn out to have an asymptotic distribution that is neither $O_p(T^{-1})$ Dickey-Fuller, nor $O_p(N^{-1/2})$ normal and asymptotically unbiased. Instead, the estimator turns out to be consistent and asymptotically normal, but has a nonvanishing bias in its asymptotic distribution.

Keywords: random field, time series, panel data, unit root

JEL Classification: C21, C22, C23

1. Introduction

It has long been known that panel data structures allow an empirical researcher to identify effects that would otherwise be unavailable using only cross-section or time-series data. Many such applications, however, really exploit only the large cross-section aspect to panels (e.g., Chamberlain, 1984; Hausman and Taylor, 1981; Holtz-Eakin, Newey, and Rosen, 1988). In fact, a number of important techniques in panel data econometric analysis quickly become intractable if the time dimension grows large. Think, for instance, of techniques that model each time period as a separate equation in a simultaneous system.

Yet, for certain economic questions and data sets, it is quite natural and appropriate to view the time dimension as being large along with that of the cross-section: For instance, in the study of macroeconomic growth and fluctuations, there is often great interest in the long-run effects of particular disturbances (e.g., Blanchard and Quah, 1989). At the same time, however, one might also be interested in the dynamic effects of those disturbances across a rich cross-section of observation units (e.g., Blanchard and Katz, 1992; Barro and Sala-i-Martin, 1991). The currently available data sets for studying the dynamics of different regions, industries, or asset prices often span time and cross-section dimensions having approximately the same order of magnitude: researchers then end up discarding useful information by pooling across time (e.g., Barro and Sala-i-Martin, 1991) or by aggregating cross sections into portfolios (e.g., Lo and McKinlay, 1988) so that they can then apply standard econometric ideas.

This paper investigates, for the canonical unit root time series regression, estimation and inference for data structures where the time-series and cross-section dimensions are comparable in magnitude. To focus on the issues that are novel in such analysis, I confine study to the simplest possible case, and use easily interpretable regularity conditions which are stronger than absolutely necessary.[1] Even so, interesting subtleties will arise: the principal result of the paper shows that the unit root regression coefficient estimator is asymptotically distributed neither (unbiased) normal at rate $O_p(N^{-1/2})$, as one might expect from standard panel data analysis, nor standard Dickey-Fuller at rate $O_p(T^{-1})$, as one might expect from standard time series analysis. Instead, the estimator is consistent and asymptotically normal, but with a nonvanishing bias in the asymptotic distribution.

The remainder of this paper is organized as follows. Section 2 calculates the asymptotic distribution of the least squares estimator for the lag coefficient when the data have a unit root in the time series dimension, and where both cross-section and time series dimensions are comparable in magnitude (i.e., the data are a *random field*). Differences from the standard time series case are described more carefully there. Section 3 reports the results of a Monte Carlo study to evaluate the accuracy of the Monte Carlo approximation. Section 4 briefly concludes; an appendix gives the proof of the main result.

[1] Quah (1992) has studied the more general case, with correspondingly more complicated proofs and manipulations. The proof for the simple case below makes the intuition in the general case transparent.

2. Asymptotic Approximation

Unit root regression for univariate time series is now well understood (Phillips, 1987). We briefly present it here only to establish notation. Suppose $\{ \epsilon(t) :$ integer $t \}$ is a mean zero random sequence satisfying a functional central limit theorem, i.e., for

$$\tilde{\mathcal{B}}_T(r) \stackrel{\text{def}}{=} T^{-1/2} \sum_{t=1}^{[rT]} \epsilon(t), \quad \text{all } r \text{ in } [0,1]$$

(by the usual convention, [ ] denotes integer part, and $\sum_{t=1}^{0} \epsilon(t)$ is taken to be zero), there exists a finite positive constant s such that:

$$s^{-1}\tilde{\mathcal{B}}_T \Rightarrow \mathcal{B} \quad \text{as } T \to \infty$$

(with $\Rightarrow$ denoting weak convergence, and $\mathcal{B}$ standard Brownian motion or Wiener process). It will be convenient in the sequel to define the normalized version $\mathcal{B}_T = s^{-1}\tilde{\mathcal{B}}_T$ so that we have $\mathcal{B}_T \Rightarrow \mathcal{B}$ as $T \to \infty$.

A weak law of large numbers for ϵ follows from the preceding, since

$$T^{-1} \sum_{t=1}^{T} \epsilon(t) = T^{-1/2} s \mathcal{B}_T(1) \xrightarrow{\text{Pr}} 0 \quad \text{as } T \to \infty \text{ (by Markov inequality)},$$

and, by application of a continuous mapping theorem (Billingsley, 1968, p. 30), so does an ordinary central limit theorem:

$$s^{-1}T^{-1/2} \sum_{t=1}^{T} \epsilon(t) = \mathcal{B}_T(1) \xrightarrow{\mathcal{L}} \mathcal{B}(1) \equiv \mathcal{N}(0,1) \quad \text{as } T \to \infty$$

(where $\equiv$ denotes equivalence in distribution).

When $\{\epsilon(t)\}$ is serially uncorrelated with variance positive and identical across t, then s simply equals $\epsilon(1)$'s standard deviation. In general, however, s is analogous to the square root of ϵ's spectral density at frequency zero,

$$\lim_{T \to \infty} \text{Var}\left(T^{-1/2} \sum_{t=1}^{T} \epsilon(t) \right).$$

We will also assume that a weak law of large numbers applies to $\{\epsilon(t)^2\}$ so that

$$T^{-1}\sum_{t=1}^{T}\epsilon(t)^2 \xrightarrow{\text{Pr}} \sigma^2 > 0 \quad \text{as } T \to \infty.$$

(This will typically follow from the same set of regularity conditions giving a functional central limit theorem.)

Suppose that the observed time series $\{\, X(t) : \text{integer } t \,\}$ is generated by:

$$X(t) = X(t-1) + \epsilon(t), \quad t \geq 1;$$

$$X(0) \text{ a given random variable.}$$

If the observed sample is $\{\, X(t) : t = 0, 1, \ldots, T \,\}$, then the least squares estimator for the regression coefficient of X on its first lag is

$$b_T = \left(\sum_{t=1}^{T} X(t-1)^2\right)^{-1}\left(\sum_{t=1}^{T} X(t)X(t-1)\right),$$

so that

$$T(b_T - 1) = \left(T^{-2}\sum_{t=1}^{T} X(t-1)^2\right)^{-1}\left(T^{-1}\sum_{t=1}^{T} X(t-1)\epsilon(t)\right).$$

Recalling that

$$X(t) = X(0) + \sum_{t=1}^{t}\epsilon(t) = X(0) + sT^{1/2}\mathcal{B}_T(t/T),$$

it is straightforward to show (e.g., Phillips, 1987):

$$T^{-1}\sum_{t=1}^{T} X(t-1)\epsilon(t) = \frac{1}{2}\left[s^2\mathcal{B}_T(1) - T^{-1}\sum_{t=1}^{T}\epsilon(t)^2\right] + o_p(1)$$

and

$$T^{-2}\sum_{t=1}^{T} X(t-1)^2 = s^2\int_0^1 \mathcal{B}_T(r)^2\,dr + o_p(1).$$

Applying the observations above, we immediately have:

Theorem 2.1: *As $T \to \infty$, the estimator b_T behaves as:*

$$T(b_T - 1) \Rightarrow \left(\int_0^1 \mathcal{B}(r)^2\,dr\right)^{-1} \cdot \frac{1}{2}\left(\mathcal{B}(1)^2 - \sigma^2/s^2\right).$$

A number of features in this result are useful to note here, for comparison with those below. First, the least squares estimator b_T converges to the correct value of unity at rate T, faster than the usual $T^{1/2}$ rate in ordinary regression. Second, the initial condition $X(0)$ is asymptotically irrelevant. Third, the approximating random variable on the left hand side bears a non-normal distribution, one that does not in general have expectation zero. Fourth, the numerator random variable is a shifted $\chi^2(1)$ with mean $1 - \sigma^2/s^2$ (zero when ϵ is serially uncorrelated, but not otherwise), while the denominator is a nondegenerate positive random variable. Nevertheless, the distribution of the ratio is easily generated by Monte Carlo simulation; its critical points have been tabulated, for instance, in Fuller (1976) Table 8.5.1.

We turn now to the situation of interest, where we have an extensive cross-section of observations

$$\{\, X_j(t) : j = 1, 2, \ldots, N;\, t = 0, 1, \ldots, T\,\},$$

which, for each j, is generated by

$$X_j(t) = X_j(t-1) + \epsilon_j(t), \quad t \geq 1;$$

$$X_j(0) \text{ a given random variable.}$$

The simplest case arises when observations are independent in the cross-section: this is standard in panel data analysis, although in time series econometrics, positing independence across observations is unusual. Modelling cross-sectional dependence is complicated considerably by the fact that, unlike in a time series, individual observations in a cross section need display no natural ordering. Thus, the interpretation of mixing conditions (say) in cross-section economic data is unclear—it is not evident what is meant by independence for observations "sufficiently far apart". One possibility for modelling dependence in dynamic cross sections might be a structure like that in Quah and Sargent (1993), although as Geweke (1993) emphasizes, a rigorous inference theory there too has yet to be developed. Yet another possibility in such data sets with rich cross-section and time-series variation is to eschew regression analysis altogether and to model the data as a dynamically evolving distribution. [Some economic models even suggest this as the natural econometric structure to investigate particular questions (see Quah, 1993a, b, c).]

Instead of the standard panel data setting where the researcher is concerned with unobservable individual effects and a fixed, finite time dimension T, here we ignore the first issue, and take N and T to be the same order of magnitude, $N = N(T) = O(T)$. We do this to focus on how this new data structure affects the time series results given above in Theorem 2.1 and its surroundings.

By analogy with the time series case, take the estimator for the regression coefficient of X on its own first lag to be:

$$b_T = \left(\sum_{j=1}^{N(T)} \sum_{t=1}^{T} X_j(t-1)^2\right)^{-1} \left(\sum_{j=1}^{N(T)} \sum_{t=1}^{T} X_j(t)X_j(t-1)\right).$$

Notice that the terms that appear on the right hand side are *not* those that would obtain by stacking the data as in, e.g., Holtz-Eakin, Newey, and Rosen (1988).

These terms are instead, when appropriately normalized, sample analogues of certain (conditional) population moments.

For random variable Y with finite p-th absolute moment, $E|Y|^p < \infty$, define p-norm as

$$\|Y\|_p = (E|Y|^p)^{1/p} = E^{1/p}|Y|^p.$$

The asymptotic distribution of b_T is then given in the following.

Theorem 2.2: *Assume that* $\{\, \epsilon_j(t) \,:\, \textit{integer } j,\, t \,\}$ *is a collection of independent random variables, and* $\{X_j(0) \,:\, \textit{integer } j\,\}$ *is a sequence of independent random variables such that:*

(i) $E\epsilon_j(t) = 0$ *and* $0 < \mathrm{Var}(\epsilon_j(t)) = \sigma^2 < \infty$ *for all* j *and* t*; and*

(ii) for all j,

$$E\left(X_j(0) \cdot \left[T^{-1/2} \sum_{t=1}^{T} \epsilon_j(t) \right] \right) \to \mu \quad \text{as } T \to \infty \text{ with } |\mu| < \infty.$$

Further, assume that for some positive number δ,

(iii) $\sup_{j,t} \|\epsilon_j(t)\|_{4+\delta} < \infty$;

(iv) $\sup_{j,T} \|T^{-1/2} \sum_{t=1}^{T} \epsilon_j(t)\|_{4+\delta} < \infty$ *and;*

(v) $\sup_j \|X_j(0)\|_{2+\delta} < \infty$.

Then, for $N = N(T) = \kappa T$ *with* $\kappa > 0$, *we have:*

$$2^{-1/2} N(T)^{1/2} T \left(b_T - 1 - 2\frac{\mu}{\sigma^2} T^{-3/2} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \quad \text{as } T \to \infty.$$

The proof of this result is given in the Appendix, but some remarks are appropriate here: notice that the convergence rate is $N(T)^{1/2}T$, or simply $T^{3/2}$, under our assumption that $N(T)$ is κT. In any application, N and T are fixed and given; thus *any* assumption we make on the relation between them as each gets large is necessarily arbitrary. I have chosen what seems to me the natural normalization. The assumption could be relaxed to be $N = O(T)$ without loss, but *some* such assumption will certainly be needed.

The resulting rate of $T^{3/2}$ can be viewed as multiplying the rate $N^{1/2}$ from standard regression with the rate T from unit root time series regression. The theorem asserts that the estimator b_T is consistent for the correct value of unity, but the asymptotic distribution has a nonzero mean of $2\mu/\sigma^2$ which depends on the covariances of the initial condition $X(0)$ with subsequent ϵ's as well as the variance of the ϵ's. Thus, unlike the time series case of Theorem 2.1, initial conditions do matter here—even as T gets arbitrarily large.

In the time series case, the numerator random variable of the asymptotic approximation has zero mean when the ϵ's are serially uncorrelated; here, however, the mean of the asymptotic distribution is nonzero even when the ϵ's are serially independent. Notice that in condition (ii), under the other assumptions, the product's second term $T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t)$ is $O_p(1)$. The moment conditions (iii)–(v) require only a little more than bounded fourth and second moments on ϵ and $X(0)$ respectively. In (iv), the term $T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t)$ is, again, seen to be just $O_p(1)$, and converges to a normal random variable; the last of course has all moments finite. While more primitive conditions might be available that would imply (iv), they would add no further insight in the current discussion. Finally, notice that if ϵ were iid normal, then (iii) and (iv) would automatically hold.

These conditions are not the weakest possible, but they are easy to verify; further, in the proof, they illustrate the reasoning giving rise to the result without unnecessary and distracting complications.

3. Monte Carlo Results

This section reports the results from a Monte Carlo study to assess the small-sample accuracy of Theorem 2.2. The Table gives the critical values for different tail probabilities from a Monte Carlo sample of 10,000 draws. The experiments here take the (nuisance) parameters μ and σ—for which it is easy to get consistent estimators—as known.

I consider 25 different settings for N and T, each ranging from 25 to 1000. Looking down the columns of Table 1 gives—for varying values of N and T—Monte Carlo critical values for different tail probabilities, the latter ranging across the rows. The last two rows also show the asymptotic critical values for the standard cross-section/panel data regression ($N = \infty$, $T = 1$) and for the standard Dickey-Fuller time series regression ($N = 1$, $T = \infty$). Thus, the last but one row simply tabulates the standard normal, while the last row reproduces Table 8.5.1 from Fuller (1976).

This table makes clear that large N and T drive the distribution of the estimator towards the normal: Small N and T give rise to the same asymmetry that describes the Dickey-Fuller $(1, \infty)$ distribution, while both the large N, small T and simultaneously large N and T cases are well-approximated by the standard normal distribution.[2]

Note that the table already corrects for the asymptotic bias, and thus large N—with both small and large T—should (and does) have the same asymptotics.

[2] The unit roots case with large N and small T had also been suggested on page 1373 of Holtz-Eakin, Newey, and Rosen (1988).

More extensive experiments have been carried out—all verifying the asymptotic approximations of the previous section and the appendix. For reasons of space, however, they are not presented here. (See Quah, 1992.)

4. Conclusion

This paper has begun analysis of the subtleties that arise in unit-roots regression in data that have simultaneously extensive cross-section and time-series variation. The asymptotic distribution derived here can be understood as a mixture of the standard normal and Dickey-Fuller-Phillips asymptotics.

Economists (macroeconomists in particular) are now considering progressively richer models where the natural datasets to study are no longer time series or standard cross-sections or panels. The analytical results in this note should serve as a useful beginning to allow more complete and rigorous econometric analysis of such situations.

References

Andrews, Donald W. K. (1988) "Laws of large numbers for dependent non-identically distributed random variables." *Econometric Theory* 4(3), 458–467, December

Barro, Robert J., and Xavier Sala-i-Martin (1991) "Convergence across states and regions." *Brookings Papers on Economic Activity* 1, 107–182, April

Billingsley, Patrick (1968) *Convergence of Probability Measures* (New York NY: John Wiley)

Blanchard, Olivier Jean, and Danny Quah (1989) "The dynamic effects of aggregate demand and supply disturbances." *American Economic Review* 79(4), 655–673, September

Blanchard, Olivier Jean, and Lawrence F. Katz (1992) "Regional evolutions." *Brookings Papers on Economic Activity* 1, 1–75, April

Chamberlain, Gary (1984) "Panel data." In *Handbook of Econometrics vol. II,* ed. Zvi Griliches and Michael D. Intriligator (Amsterdam: Elsevier North-Holland) chapter 22, pp. 1247–1318

Fuller, Wayne A. (1976) *Introduction to Statistical Time Series* (New York NY: John Wiley)

Geweke, John (1993) "Comments on Quah and Sargent: A dynamic index model for large cross sections." In *New Research on Business Cycles, Indicators, and Forecasting,* ed. James Stock and Mark Watson (Chicago IL: University of Chicago Press)

Hausman, Jerry A., and William E. Taylor (1981) "Panel data and unobservable individual effects." *Econometrica* 49(6), 1377–1398, November

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen (1988) "Estimating vector autoregressions with panel data." *Econometrica* 56(6), 1371–1395, November

Lo, Andrew W., and A. Craig McKinlay (1988) "Stock market prices do not follow random walks: Evidence from a simple specification test." *Review of Financial Studies* 1(1), 41–66, Spring

Phillips, Peter C. B. (1987) "Time series regression with unit roots." *Econometrica* 55(2), 277–302, March

Quah, Danny (1992) "International patterns of growth: I. Persistence in cross-country disparities." Working Paper, LSE, London, October

__________ (1993a) "Dependence in growth and fluctuations across economies with mobile capital." Working Paper, Economics Department, LSE, London WC2A 2AE, July

__________ (1993b) "Empirical cross-section dynamics in economic growth." *European Economic Review* 37, 426–434, April

__________ (1993c) "Galton's fallacy and tests of the convergence hypothesis." *The Scandinavian Journal of Economics* 95(4), 427–443, December

Quah, Danny, and Thomas J. Sargent (1993) "A dynamic index model for large cross sections." In *Business Cycles, Indicators, and Forecasting*, ed. James Stock and Mark Watson, vol. 28 (Chicago IL: University of Chicago Press and NBER) chapter 7, pp. 285–306

Wooldridge, Jeffrey M., and Halbert White (1988) "Some invariance principles and central limit theorems for dependent heterogenous processes." *Econometric Theory* 4(2), 210–230, August

5. Appendix

This appendix contains the proof of the Theorem in the paper.

Proof of Theorem 2.2: Define for each j the Brownian motion approximant

$$\mathcal{B}_{jT}(r) \stackrel{\text{def}}{=} \sigma^{-1}T^{-1/2}\sum_{t=1}^{[rT]}\epsilon_j(t), \quad \text{for } r \text{ in } [0,1]$$

(recognizing that $s = \sigma$ when, for fixed j, the sequence $\{\epsilon_j(t)\}$ comprises uncorrelated random variables). Note that (ii) implies

$$\forall j : E(X_j(0)\mathcal{B}_{jT}(1)) \to \sigma^{-1}\mu \quad \text{as } T \to \infty.$$

From the definition of b_T we have:

$$b_T - 1 - 2\frac{\mu}{\sigma^2}T^{-3/2} = \left(\sum_j\sum_{t=1}^{T}X_j(t-1)^2\right)^{-1}$$
$$\times\left(\sum_j\sum_{t=1}^{T}X_j(t-1)\epsilon_j(t) - 2\frac{\mu}{\sigma^2}T^{-3/2}\sum_j\sum_{t=1}^{T}X_j(t-1)^2\right).$$

Take the denominator: performing the usual time series calculations for each j gives

$$\sum_j\sum_{t=1}^{T}X_j(t-1)^2 = T\sum_j X_j(0)^2$$
$$+ 2\sigma T^{3/2}\sum_j X_j(0)\left[\int_0^1 \mathcal{B}_{jT}(r)\,dr - T^{-1}\mathcal{B}_{jT}(1)\right]$$
$$+ \sigma^2T^2\sum_j\left[\int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2\right].$$

Normalizing by T^2N, this obeys

$$(T^2N(T))^{-1}\sum_{j=1}^{N(T)}\sum_{t=1}^{T}X_j(t-1)^2 \xrightarrow{\text{Pr}} \sigma^2/2 \quad \text{as } T \to \infty. \tag{5.1}$$

To see this, consider each of the summands in turn. First,

$$(T^2N(T))^{-1}T\sum_{j=1}^{N(T)} X_j(0)^2 = T^{-1}N(T)^{-1}\sum_{j=1}^{N(T)} X_j(0)^2 \xrightarrow{\text{Pr}} 0 \quad \text{as } T\to\infty,$$

from Markov inequality combined with

$$\begin{aligned}\left\|T^{-1}N^{-1}\sum_{j=1}^{N} X_j(0)^2\right\|_1 &\le T^{-1}N^{-1}\sum_{j=1}^{N}\|X_j(0)^2\|_1\\ &\le T^{-1}\sup_j\|X_j(0)\|_2^2 \to 0 \quad \text{as } T\to\infty\end{aligned}$$

given (v) (using Liapounov inequality). Next, we show that

$$T^{-1/2}N(T)^{-1}\sum_{j=1}^{N(T)} X_j(0)\left[\int_0^1 \mathcal{B}_{jT}(r)\,dr - T^{-1}\mathcal{B}_{jT}(1)\right] \xrightarrow{\text{Pr}} 0. \tag{5.2}$$

This follows from:

$$\begin{aligned}\int_0^1 \mathcal{B}_{jT}(r)\,dr - T^{-1}\mathcal{B}_{jT}(1) &= \sum_{t=1}^{T}\mathcal{B}_{jT}((t-1)/T)\cdot T^{-1}\\ &= \sigma^{-1}T^{-3/2}\sum_{t=1}^{T}\left(\sum_{l=1}^{t-1}\epsilon_j(l)\right)\\ &= \sigma^{-1}T^{-1/2}\sum_{t=1}^{T}(1-t/T)\epsilon_j(t),\end{aligned}$$

so that

$$\begin{aligned}&T^{-1/2}N(T)^{-1}\sum_{j=1}^{N(T)} X_j(0)\left[\int_0^1 \mathcal{B}_{jT}(r)\,dr - T^{-1}\mathcal{B}_{jT}(1)\right]\\ &\qquad = \sigma^{-1}T^{-1/2}N(T)^{-1}\sum_{j=1}^{N(T)} X_j(0)\left(T^{-1/2}\sum_{t=1}^{T}(1-t/T)\epsilon_j(t)\right).\end{aligned}$$

But

$$\left\| N^{-1}\sum_{j=1}^{N} X_j(0)\left(T^{-1/2}\sum_{t=1}^{T}(1-t/T)\epsilon_j(t)\right)\right\|_1$$
$$\leq N^{-1}\sum_{j=1}^{N}\|X_j(0)\|_2\cdot\left\|T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t)\right\|_2$$
$$\leq \sup_j\|X_j(0)\|_2\cdot\sup_{j,T}\left\|T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t)\right\|_2$$

by the Minkowski and Hölder inequalities. From (iv) and (v) and the Liapounov inequality, the right hand side above is finite independent of j and T; combined with the Markov inequality, this establishes (5.2). Finally, it only remains to verify

$$N(T)^{-1}\sum_{j=1}^{N(T)}\left[\int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2\right]\xrightarrow{\Pr}\frac{1}{2}\quad \text{as } T\to\infty. \tag{5.3}$$

Notice that for all T, the individual summands $\int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2$ are independent across j. Expanding each summand,

$$\int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2 = \sum_{t=1}^{T}T^{-1}\mathcal{B}_{jT}((t-1)/T)^2$$
$$= \sigma^{-2}T^{-1}\sum_{t=1}^{T}\left(T^{-1/2}\sum_{l=1}^{t-1}\epsilon_j(l)\right)^2,$$

so that the expectation of each satisfies:

$$E\left[\int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2\right] = \sigma^{-2}T^{-2}\sum_{t=1}^{T}(t-1)\sigma^2$$
$$=\frac{1}{2}(1-T^{-1})\to\frac{1}{2}\quad\text{as } T\to\infty$$

uniformly in j, using the uncorrelatedness of $\{\,\epsilon_j(t)\ :\ t\,\}$. Further, there exists some positive δ such that:

$$\begin{aligned}\left\| \int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2 \right\|_{1+\delta} &\le \sigma^{-2}T^{-1}\sum_{t=1}^{T}\left\| \left(T^{-1/2}\sum_{l=1}^{t-1}\epsilon_j(l)\right)^2 \right\|_{1+\delta} \\ &\le \sigma^{-2}T^{-1}\sum_{t=1}^{T}\left\| \left(T^{-1/2}\sum_{l=1}^{t-1}\epsilon_j(l)\right) \right\|_{2(1+\delta)}^2 \\ &\le \sigma^{-2}\sup_{j,T}\left\| T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t) \right\|_{2(1+\delta)}^2 \\ &< \infty\end{aligned}$$

by the Minkowski and Liapounov inequalities and assumption (iv). Thus the family $\left\{ \int_0^1 \mathcal{B}_{jT}(r)^2\,dr - T^{-1}\mathcal{B}_{jT}(1)^2 \right\}$ is uniformly integrable in j and T (e.g., Billingsley 1968, p. 32). The result (5.3) then follows by a weak law of large numbers (Andrews 1988, p. 462, item 1).

Turn next to the numerator. Normalized by $TN^{1/2}$, this is:

$$\begin{aligned}T^{-1}N^{-1/2}\sum_j\sum_t X_j(t-1)\epsilon_j(t) - 2\frac{\mu}{\sigma^2}T^{-5/2}N^{-1/2}\sum_j\sum_t X_j(t-1)^2 \\ = T^{-1/2}N^{-1/2}\sum_j X_j(0)\,(\sigma\mathcal{B}_{jT}(1)) \\ + \frac{1}{2}N^{-1/2}\left[\sigma^2\sum_j \mathcal{B}_{jT}(1)^2 - T^{-1}\sum_j\sum_t \epsilon_j(t)^2\right] \\ - 2\frac{\mu}{\sigma^2}T^{-1/2}N^{1/2}\left[(T^2N)^{-1}\sum_j\sum_t X_j(t-1)^2\right].\end{aligned}$$

Adding and subtracting the term $-\mu T^{-1/2}N^{1/2}$, this is

$$\begin{aligned}T^{-1/2}N^{-1/2}\sum_j\left[X_j(0)\,(\sigma\mathcal{B}_{jT}(1)) - \mu\right] \\ + \frac{1}{2}N^{-1/2}\left[\sigma^2\sum_j \mathcal{B}_{jT}(1)^2 - T^{-1}\sum_j\sum_t \epsilon_j(t)^2\right] \\ - 2\frac{\mu}{\sigma^2}T^{-1/2}N^{1/2}\left[(T^2N)^{-1}\sum_j\sum_t X_j(t-1)^2 - \frac{1}{2}\sigma^2\right].\end{aligned}$$

But from $T^{-1/2}N^{1/2} = \kappa^{1/2}$, and the previous convergence result (5.1) for the denominator, the last term is $o_p(1)$. Further, the first term too is $o_p(1)$ from

$$T^{-1/2}N(T)^{-1/2}\sum_{j=1}^{N(T)}\left[X_j(0)\left(\sigma\mathcal{B}_{jT}(1)\right)-\mu\right]$$
$$= \kappa^{-1/2}N(T)^{-1}\sum_{j=1}^{N(T)}\left[X_j(0)\left(\sigma\mathcal{B}_{jT}(1)\right)-\mu\right],$$

with the individual summands being independent and uniformly integrable, and the limiting relation $\lim_{T\to\infty} E\left[X_j(0)\left(\sigma\mathcal{B}_{jT}(1)\right)-\mu\right] = 0$. To see uniform integrability, calculate for positive δ,

$$\begin{aligned}
&E\left|X_j(0)\left(\sigma\mathcal{B}_{jT}(1)\right)\right|^{1+\delta}\\
&\quad\le E\left(|X_j(0)|^{1+\delta}|\sigma\mathcal{B}_{jT}(1)|^{1+\delta}\right)\\
&\quad= \left\|\,|X_j(0)|^{1+\delta}|\sigma\mathcal{B}_{jT}(1)|^{1+\delta}\,\right\|_1\\
&\quad\le \|X_j(0)\|_{2+2\delta}^{1+\delta}\cdot\left\|T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t)\right\|_{2+2\delta}^{1+\delta} \qquad \text{(by Hölder inequality)}\\
&\quad\le \left(\sup_j \|X_j(0)\|_{2+2\delta}\right)^{1+\delta}\left(\sup_{j,T}\left\|T^{-1/2}\sum_{t=1}^{T}\epsilon_j(t)\right\|_{2+2\delta}\right)^{1+\delta}\\
&\quad< \infty \quad \text{independent of } j \text{ and } T.
\end{aligned}$$

Thus, the numerator after normalization is asymptotically equivalent to

$$\frac{1}{2}N^{-1/2}\sum_j\left[\left(\sigma\mathcal{B}_{jT}(1)\right)^2 - T^{-1}\sum_t \epsilon_j(t)^2\right]$$
$$= N^{-1/2}\sum_j\left[T^{-1}\sum_{m=1}^{T-1}\sum_{l=m+1}^{T}\epsilon_j(l)\epsilon_j(l-m)\right].$$

Each summand is independent across j, and by the serial uncorrelatedness in ϵ_j, has expectation zero. Further, there is a positive δ such that:

$$\begin{aligned}\left\| (\sigma \mathcal{B}_{jT}(1))^2 - T^{-1} \sum_t \epsilon_j(t)^2 \right\|_{2+\delta} &\\ &\leq \sup_{j,T} \left\| T^{-1/2} \sum_{t=1}^{T} \epsilon_j(t) \right\|^2_{2(2+\delta)} + \sup_{j,T} \|\epsilon_j(t)\|^2_{2(2+\delta)} \\ &< \infty \quad \text{independent of } j \text{ and } T \text{ by (iii) and (iv).}\end{aligned}$$

Finally, it is straightforward to calculate:

$$\begin{aligned}\mathrm{Var}\left(T^{-1} \sum_{m=1}^{T-1} \sum_{l=m+1}^{T} \epsilon_j(l)\epsilon_j(l-m) \right) &= T^{-2}\frac{1}{2}T(T-1)\sigma^4 \\ &\to \frac{1}{2}\sigma^4 \quad \text{as } T \to \infty.\end{aligned}$$

Consequently, by Wooldridge and White (1988, 3.1, p.219), the normalized numerator converges in distribution to $\mathcal{N}(0, \frac{1}{2}\sigma^4)$. Recalling that the normalized denominator converges in probability to $\frac{1}{2}\sigma^2$, we have

$$2^{-1/2}N(T)^{1/2}T\left(b_T - 1 - 2\frac{\mu}{\sigma^2}T^{-3/2}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1),$$

as was to be shown. *Q.E.D.*

Table: Monte Carlo CDF: 10,000 draws (Known μ, σ)

$$2^{-1/2}N(T)^{1/2}T\left(b_T - 1 - 2\frac{\mu}{\sigma^2}T^{-3/2}\right)$$

	Probability no greater than:							
(N,T)	1%	2.5%	5%	10%	90%	95%	97.5%	99%
(25, 25)	−3.13	−2.60	−2.14	−1.64	1.11	1.42	1.66	1.93
(25, 50)	−3.19	−2.60	−2.12	−1.60	1.10	1.40	1.67	1.94
(25, 100)	−3.17	−2.57	−2.06	−1.58	1.08	1.38	1.65	1.95
(25, 250)	−3.19	−2.62	−2.14	−1.60	1.07	1.36	1.58	1.85
(25, 1000)	−3.25	−2.62	−2.13	−1.62	1.08	1.38	1.65	1.92
(50, 25)	−2.87	−2.46	−2.00	−1.55	1.16	1.47	1.76	2.06
(50, 50)	−2.86	−2.36	−1.93	−1.49	1.18	1.50	1.75	2.02
(50, 100)	−2.96	−2.37	−1.91	−1.48	1.14	1.44	1.70	2.01
(50, 250)	−2.83	−2.36	−1.93	−1.47	1.12	1.43	1.69	1.99
(50, 1000)	−2.88	−2.37	−1.97	−1.51	1.14	1.47	1.75	2.04
(100, 25)	−2.74	−2.29	−1.90	−1.47	1.20	1.51	1.80	2.13
(100, 50)	−2.65	−2.20	−1.82	−1.39	1.22	1.57	1.89	2.17
(100, 100)	−2.56	−2.15	−1.81	−1.39	1.19	1.51	1.80	2.12
(100, 250)	−2.65	−2.20	−1.81	−1.39	1.18	1.47	1.77	2.08
(100, 1000)	−2.68	−2.22	−1.83	−1.40	1.18	1.50	1.80	2.13
(250, 25)	−2.61	−2.17	−1.85	−1.43	1.24	1.57	1.87	2.23
(250, 50)	−2.45	−2.06	−1.73	−1.31	1.29	1.67	1.96	2.31
(250, 100)	−2.46	−2.05	−1.70	−1.31	1.24	1.58	1.90	2.23
(250, 250)	−2.44	−2.08	−1.75	−1.34	1.23	1.55	1.83	2.16
(250, 1000)	−2.54	−2.07	−1.73	−1.33	1.19	1.54	1.84	2.16
(1000, 25)	−2.44	−2.04	−1.74	−1.36	1.27	1.61	1.89	2.24
(1000, 50)	−2.36	−1.94	−1.59	−1.20	1.37	1.74	2.02	2.37
(1000, 100)	−2.39	−1.94	−1.63	−1.25	1.33	1.69	1.95	2.26
(1000, 250)	−2.44	−2.02	−1.70	−1.31	1.28	1.62	1.91	2.26
(1000, 1000)	−2.39	−2.04	−1.73	−1.31	1.23	1.58	1.85	2.18
$(\infty, 1)$	−2.33	−1.96	−1.64	−1.29	1.29	1.64	1.96	2.33
$(1, \infty)$	−13.8	−10.5	−8.1	−5.7	0.93	1.28	1.60	2.03

FINANCIAL MARKETS GROUP - DISCUSSION PAPER SERIES

133 Webb, David, "Project Selection with Screened and Contingent Debt", (Feb'92)

134 Sentana, Enrique, Mushtaq Shah and Sushil Wadhwani, "Has the EMS Reduced the Cost of Capital?", (Mar'92)

135 Sentana, Enrique, "Factor Representing Portfolios in Large Asset Markets", (Mar'92)

136 Connor, Gregory and Richard Breen, "Non-Arbitrage and Recursive Competitive Equilibrium Pricing", (May'92).

137 Connor, Gregory and Robert A. Korajczyk, "A Test for the Number of Factors in an Approximate Factor Model", (May'92).

138 Hansen, Eric, "The Role of Asymmetric Information in Project Financing Decisions", (May'92).

139 Sentana, Enrique, "Identification of Multivariate Conditionally Heteroskedastic Factor Models", (May'92).

140 Demos, Antonis and Enrique Sentana, "An EM-Based Algorithm for Conditionally Heteroskedastic Factor Models", (May'92).

141 Evans, George W. and Garey Ramey, "Expectation Calculation, Hyperinflation and Currency Collapse", (June'92).

142 Curcio, Riccardo and C.A.E. Goodhart, "When Support/Resistance Levels are Broken, Can Profits be Made? Evidence from the Foreign Exchange Market", (July'92).

143 Vassilicos, J.C., Antonis Demos and Fidelio Tata, "No Evidence of Chaos But Some Evidence of Multifractals in the Foreign Exchange and the Stock Markets", (June'92).

144 Perotti, Enrico C., and Kathryn E. Spier, "Capital Structure as a Bargaining Tool: The Role of High Leverage in Contract Renegotiation", (July 1992).

145 Kulatilaka, Nalin and Enrico C. Perotti, "Strategic Investment Timing Under Uncertainty", (July 1992).

146 Jeanne, Olivier, "Monetary Policy in England 1893-1914: A Structural VAR Analysis", (July 1992).

147 Gale, Douglas, "Informational Capacity and Financial Collapse", (July 1992).

148 Aghion, Philippe, Oliver Hart and John Moore, "The Economics of Bankruptcy Reform", (July 1992).

149 Connor, Gregory and Robert A. Korajczyk, "The Arbitrage Pricing Theory and Multifactor Models of Asset Returns", (October 1992).

150 Pagano, Marco and Ailsa Röell, "Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading", (October 1992).

151 Evans, George W. and Seppo Honkapohja, "Learning and Economic Fluctuations: Using Fiscal Policy to Steer Expectations", (October 1992).

152 Demos, Antonis and Charles A.E. Goodhart, "The Interaction between the Frequency of Market Quotations, Spread, and Volatility in the Foreign Exchange Market", (October 1992).

153 Pagano, Marco, "Financial Markets and Growth: An Overview", (October 1992).

154 Quah, Danny, "Empirical Cross-Section Dynamics in Economic Growth", (November 1992).

155 Bolton, Patrick and Gérard Roland, "The Economics of Mass Privatization: Czechoslovakia, East Germany, Hungary and Poland", (November 1992).

156 Chen, Zhaohui, "Cointegration and Exchange Rate Forecasting: A State Space Model, (February 1993).

157 Hart, Oliver, "An Economist's View of Fiduciary Duty", (February 1993).

158 Breedon, F.J., "Intraday Price Formation on the London Stock Exchange", (May 1993)

159 Davis, E.P., "VAR Modelling of the German Economy with Financial Spreads as Key Indicator Variables", (May 1993).

160 Gregory, Alan, John Matatko, Ian Tonks and Richard Purkis, "UK Directors' Trading: The Impact of Dealings in Smaller Firms", (May 1993).

161 Bolton, Patrick and David S. Scharfstein, "Optimal Debt Structure with Multiple Creditors", (July 1993).

162 Demos, Antonis, Enrique Sentana & Mushtaq Shah, "Risk and Return in January: Some UK Evidence", (July 1993).

163 Flood, Robert P. and Andrew K. Rose, "Fixing Exchange Rates: A Virtual Quest for Fundamentals", (July 1993).

164 Rose, Andrew K. and Lars E.O. Svensson, "European Exchange Rate Credibility Before the Fall", (July 1993).

165 Allen, Franklin and Douglas Gale, "Limited Market Participation and Volatility of Asset Prices", (July 1993).

166 Harvey, Andrew C. and Neil Shephard, "The Econometrics of Stochastic Volatility", (July 1993).

167 Aghion, Philippe, Oliver Hart and John Moore, "A Proposal for Bankruptcy Reform in the U.K.", (August 1993).

168 Hart, Oliver and John Moore, "Debt and Seniority: An Analysis of the Role of Hard Claims in Constraining Management", (August 1993).

169 de Jong, Frank, Theo Nijman and Ailsa Röell, "A Comparison of the Cost of Trading French Shares on the Paris Bourse and on SEAQ International", (August 1993).

170 Aghion, Philippe and Patrick Bolton, "A Theory of Trickle-Down Growth and Development with Debt-Overhang", (October 1993).

Subject to availability, copies of these Discussion Papers can be obtained from the Financial Markets Group (Room R.511, ext. 7002).