SPECTRAL MAXIMUM LIKELIHOOD ESTIMATION OF A SIGNAL TO NOISE RATIO LYING IN THE VICINITY OF ZERO

F. J. Fernández-Macho
Dpto. de Econometría y Estadística and Instituto de Economía Pública
Universidad del País Vasco-Euskal Herriko Unibertsitatea
Facultad de Ciencias Económicas
Lehendakari Aguirre 83, E48015, Bilbao, Spain

e-mail: `jf@alcib.bs.ehu.es`

as published in *Journal of Time Series Analysis*, vol. 17,
Blackwell: 1996; pp. 447-459

Abstract

A time-series model representing a decomposition into permanent plus transient components contains a deterministic component when the signal-to-noise ratio is equal to zero; otherwise, the permanent component is said to be stochastic. This distinction has important consequences in the analysis of economic phenomena. On the other hand, the absence of a stochastic permanent component in residuals from a time-series regression may indicate co-integration. This paper considers the frequency domain estimation of the signal-to-noise ratio in a representative of the unobserved components model class. The sampling properties of the estimator from the resulting approximate spectral likelihood differ from those observed in the time domain and they vary substantially depending on whether the overall slope must be estimated or not. Further it is shown that spectral estimates are T-consistent —instead of T^2-consistent in the time domain. These results may explain some of the differences in estimators from frequency domain aproximations to the likelihood and exact maximum likelihood estimators, and may be of use when testing for deterministic trends.

Key Words: Boundary Estimate, Deterministic Component, Frequency Domain, Noninvertible Moving Average, Structural Time Series Model, Unit Root, Unobserved Component.

1 Introduction

Many time series, especially in macroeconomics, are usually decomposed into a secular or permanent component (signal) and a transient component (noise). It is then the case that extraction of the permanent component is one of the first steps in the analysis of such series. However isolating the underlying signal rests on certain assumptions, concerning its deterministic vs. stochastic nature, which have important consequences (see *e.g.* Nelson & Plosser 1982). In practice, the deterministic or stochastic nature of the permanent component will be taken into account depending on whether the ratio of innovation variances (signal-to-noise ratio) is estimated to be zero or not. Besides, if the series under scrutiny has been formed by the residuals from a regression on integrated variables —*e.g.* many macroeconomic variables— then the absence of a stochastic permanent component will indicate that the variables are co-integrated. Estimating a zero (or near zero) residual signal-to-noise ratio may thus be taken as an indication of co-integration which might be worth considering along with the usual tests on non-co-integration (*c.f.* Engle & Granger 1987).

This paper considers the estimation of a signal-to-noise ratio via maximization of the (spectral) likelihood function which is obtained by transforming the observations into the frequency domain (this approach has attractive advantages in practice; see Nerlove *et al* 1979 for a general discussion of frequency domain estimation.) To put matters into perspective, it investigates the properties of the spectral maximum likelihood estimator (mle) of q in a simple model of the type

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t; \qquad t = 0, 1, \ldots, T, \\ \Delta\mu_t &= \delta + \eta_t, \\ \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} &\sim \text{NID}\left[0 \quad , \quad \begin{pmatrix} \sigma^2 & 0 \\ 0 & q\sigma^2 \end{pmatrix} \right] \end{aligned} \tag{1}$$

where y_t and μ_t denote respectively the observed variable and its unobserved trend level at time t (the latter governed by a random-walk mechanism with drift δ) and ε_t and η_t are serially independent normal disturbances.

Model (1) may be relevant in itself in a number of circumstances, for instance if fitted to an observed non-seasonal time series it may help in testing for a deterministic (vs. stochastic) trend component or if fitted to residuals from a time-series regression as in Fernandez (1993) —probably with δ set to zero— the estimated value of q can help in deciding whether the regressed variables are co-integrated. Nevertheless, the simplicity of the model object of study is chiefly for convenience and extending the results to more sophisticated models including higher order trends, seasonals, &c. (*e.g.* the "basic structural" model of Harvey & Durbin 1986) is, albeit tedious, straightforward (in fact even multivariate systems, *e.g.* Fernandez & Harvey 1990, can be easily handled in a similar fashion.) It is in these more complex models where the computational advantages of the spectral likelihood are, in practice, essential *e.g.* Fernandez (1986,90)[1].

The mle of q may be expected to exhibit unusual behaviour when true q is near zero; unusual behaviour which may be considered of a similar kind as that observed in Sargan & Bhargava (1983) for MA roots near the unit circle. In a recent paper Shephard & Harvey (1990) investigate the probabilities of obtaining zero estimates in the time domain. However, as shown in what follows, the distributional aspects of using the spectral ikelihood may be quite different.

Section 2 considers the maximization of the likelihood in the frequency domain which rests on certain simplifying circularity conditions. The circularity assumption, although useful for constructing the likelihood, is seldom the case in real-world socioeconomic time series, but it can be justified asymptotically (Fuller 1976). Section 3 compares the sampling distributions of estimators obtained in the frequency domain with those obtained in the time domain. This comparison is very important because in the frequency domain we get approximate mle's whose properties this paper seeks to clarify.

The main results can be summaried as follows:

When δ is known on *a priory* grounds obtaining a local maximum of the spectral likelihood at the zero boundary turns out to be always zero independently of the true value of q.

When δ needs to be estimated it so happens that the zero frequency is dropped completely from the likelihood and, in consequence, there is now a possibility of obtaining zero boundary estimates regardless of the true value of q. It is interesting to note how when the true signal/noise ratio is actually zero such probability of a zero estimate decreases towards zero as the sample size increases —unlike time-domain analogues in Sargan & Bhargava (1983) and Shephard & Harvey (1990). In turn, obtaining deterministic components —when they actually are not— will be less probable in the frequency domain. In this respect the differences between both distributions are noted and quantiles of the mle distribution are produced which may help in testing the deterministic vs. stochastic nature of a trend component.

With respect to the rate of consistency towards a true value $q = 0$, the spectral mle is shown to be in practice much slower than $O_p(T^{-2})$ in time domain counterparts: actually it is of $O_p(T^{-1})$ for non-circular processes.

[1]Even in univariate applications frequency domain estimators are handled more often. Thus, for example, it is the default option in STAMP —the standard computer package for the analysis and modelling of Structural Time Series Models; see J. of Applied Econometrics, vol.4 (1989) p.195.

2 Maximising the Likelihood in theFrequency Domain

Circularity means that observations are obtained as if drawn from around a circle. The circularity assumption provides a very convenient approach to obtaining the properties of estimators in the frequency domain. Indeed if (weakly) stationary processes are in general of a non-circular nature then the respective (Toeplitz) covariance matrices are asymptotically, as opposed to exactly, diagonalized by the Fourier transform (*e.g.* Fuller 1976). Thus we will still be able to use an asymptotic approximation to the likelihood which is much simpler to handle than the time-domain exact likelihood (*e.g.* Harvey & Peters 1990 or Fernandez 1990) and, in consequence, the properties of the estimators can easily be derived.

In a circular data generating process —such that T observations are somehow obtained as if drawn from around a circle— we have an autocovariance function satisfying $\gamma(\tau) = \gamma(T - \tau)$. If, besides, the process is (weakly) stationary the covariance matrix is a Toeplitz circulant which, as it is well known (*e.g.* Fuller 1976), is exactly diagonalized by the Fourier transform. This result allows us to handle the likelihood most easily by transforming the observations into the frequency domain.

Taking first differences in (1) we get the stochastic process

$$z_t = \Delta y_t - \delta = \eta_t + \Delta\varepsilon_t; \qquad t = 1, \ldots, T, \tag{2}$$

for which the circularity condition would imply

$$\varepsilon_T = \varepsilon_0. \tag{3}$$

By denoting realizations of $\{z_t\}$ and $\{y_t\}$ as $\mathbf{z} = (z_1, z_2, \ldots, z_T)'$ and $\mathbf{y} = (y_0, y_1, y_2, \ldots, y_T)'$, respectively and similarly $\mathbf{h} = (\eta_1, \eta_2, \ldots, \eta_T)'$, $\mathbf{e} = (\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T)'$, and $\mathbf{i} = (1, 1, \ldots, 1)'$, we may write (2) in a more convenient form as

$$\mathbf{z} = \mathbf{B}\mathbf{y} - \delta\mathbf{i} = \mathbf{h} + \mathbf{B}\mathbf{e},$$

where $\mathbf{B}$ is the $T \times (T + 1)$ matrix whose (i, j)-th element is

$$b_{ij} = \begin{cases} -1, & i = j, \\ 1, & i = j - 1, \\ 0, & \text{otherwise.} \end{cases}$$

Taking circularity (*i.e.* (3)) into account the last column of $\mathbf{B}$ merges into the first and the joint distribution law of $\mathbf{z}$ can be deduced to be

$$\mathbf{z} \sim \mathrm{N}\left[0, \sigma^2(q\mathbf{I}_T + \mathbf{A})\right], \tag{4}$$

where $\mathbf{A}$ is the $T \times T$ matrix whose (i, j)-th element is (note the nonzero corner elements induced by circularity)

$$a_{ij} = \begin{cases} 2, & i = j, \\ -1, & i = j \pm 1, \\ -1, & i = j \pm (T-1), \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathbf{F}$ be the $T \times T$ orthogonal Fourier matrix containing $T^{-1/2} \exp\{-\mathrm{i}2\pi(k-1)t/T\}$ as its (k,t)-th element and let $\bar{\mathbf{F}}$ denote its complex conjugate. Matrix $\mathbf{F}$ will diagonalize $\mathbf{A}$ to give the $T \times T$ diagonal matrix $\mathbf{C} = \mathbf{F}\mathbf{A}\bar{\mathbf{F}}'$ whose diagonal elements are 2π times values of the (pseudo)-spectral density of a random walk process at frequencies $\lambda_j = 2\pi j/T$, $j = 0, 1, \ldots, T-1$. By transforming $\mathbf{z}$ through rows of $\mathbf{F}$ we then obtain a complex-valued process $\{w_j = T^{-1/2} \sum_t \mathrm{e}^{-\mathrm{i}\lambda_j t} z_t\}$ whose distribution can be deduced as

$$w_j \sim \mathrm{NID}_{\mathcal{C}} \left[0, \sigma^2(q + c_j)\right],$$

$$c_0 = 0; \qquad c_j = c_{T-j} = 4\sin^2(\lambda_j/2).$$

In consequence the log-likelihood function may be written as a summation of independent terms. After concentrating σ^2 out we may express its mle and the concentrated log-likelihood as

$$\tilde{\sigma}^2(q) = \frac{1}{T} \sum_{j=0}^{T-1} \frac{1}{q + c_j} 2\pi P_j,$$

$$\mathcal{L}(q) \propto -\log\left(\frac{1}{T} \sum_{j=0}^{T-1} \frac{1}{q + c_j} 2\pi P_j\right) - \frac{1}{T} \sum_{j=0}^{T-1} \log(q + c_j), \tag{5}$$

with $P_j = P_{T-j} = (2\pi)^{-1}|w_j|^2$, $j = 0, 1, \ldots, \lfloor T/2 \rfloor$, being periodogram ordinates of the $\{z_t\}$ process, and its derivative or score function is proportional to

$$\mathcal{S}_1(\bar{q}, q) = \left(\frac{1}{T} \sum_{j=0}^{T-1} \frac{1}{q + c_j} 2\pi P_j\right)^{-1} \left(\frac{1}{T} \sum_{j=0}^{T-1} \frac{1}{(q + c_j)^2} 2\pi P_j\right) - \frac{1}{T} \sum_{j=0}^{T-1} \frac{1}{q + c_j}. \tag{6}$$

Expression (5) provides an alternative to the construction of the likelihood via the Kalman filter and the prediction-error decomposition which is easily generalizable to more complex models —see *e.g.* the multivariate case in Fernandez (1990).

We may note that, since $\{z_t\}$ depends on the value of δ, its periodogram cannot be directly computed from the data unless δ is assumed to be known *a priory* —usually $\delta = 0$, *e.g.* the local level model in Shephard & Harvey (1990)— or else it is estimated. We shall assume for the moment that δ is known and leave the handling of its mle until later.

Without the circularity condition the joint distribution law of $\mathbf{z}$ is still written as (4) except for matrix $\mathbf{A}$, whose nonzero corner elements are dropped. Without them the Fourier matrix diagonalizes $\mathbf{A}$ asymptotically only but the spectral approximation to the likelihood function may still be written as a summation of independent terms. However, the fact that in finite samples the computed periodogram ordinates are approximations to the P_j's appearing in the expressions given above must be taken into account when considering probability limits and the consistency rate of the estimator. It is known that the periodogram ordinates differ (on average) by a term of $O(T^{-1})$; more explicitly it can be shown that

$$w_j \sim \mathrm{N}_{\mathcal{C}}[0, \sigma^2(q + c_j + e_j)], \qquad \text{where} \quad e_j = f_{j1}\bar{f}_{jT} + \bar{f}_{j1}f_{jT} = \frac{2}{T}\cos\lambda_j$$

(f_{jt} denoting the (j,t)-th element of the Fourier matrix $\mathbf{F}$). Therefore the score (6) may now be expressed as proportional to

$$\mathcal{S}_1(\bar{q}, q) = \frac{\beta_1(\bar{q}, q)}{\alpha_1(\bar{q}, q)} - \gamma_1(q), \tag{7}$$

$$\begin{aligned}
\alpha_1(\bar{q}, q) &= \frac{1}{T}\sum_{j=0}^{T-1} \frac{\bar{q} + c_j + e_j}{q + c_j}\chi_j^2, \quad \bar{q} > 0; \qquad \alpha_1(0, q) = \frac{1}{T}\sum_{j=1}^{T-1} \frac{c_j + e_j}{q + c_j}\chi_j^2, \\
\beta_1(\bar{q}, q) &= \frac{1}{T}\sum_{j=0}^{T-1} \frac{\bar{q} + c_j + e_j}{(q + c_j)^2}\chi_j^2, \quad \bar{q} > 0; \qquad \beta_1(0, q) = \frac{1}{T}\sum_{j=1}^{T-1} \frac{c_j + e_j}{(q + c_j)^2}\chi_j^2, \\
\gamma_1(q) &= \frac{1}{T}\sum_{j=0}^{T-1} \frac{1}{q + c_j};
\end{aligned}$$

expression (7) being obtained after taking into account that $w_j \sim \mathrm{NID}_{\mathcal{C}}[0, \sigma^2(\bar{q}+c_j+e_j)]$ for a given true value $\bar{q}$ and that $\chi_j^2 \equiv \chi_{T-j}^2$, for $j = 0, 1, \ldots, \lfloor T/2 \rfloor$, represent mutually independent variates each with an identical chi-squared distribution of 2 degrees of freedom —1 degree of freedom if $j = 0$ or $j = \lfloor T/2 \rfloor$ and T is even.

Now for a local maximum of $\mathcal{L}(q)$ to occur at the zero boundary it is necessary that the score be negative at values of q approaching zero (from above). Since $c_0 = 0$, $\mathcal{S}_1(\bar{q}, q)$ goes to $+\infty$ as $q \to 0$, and this regardless of whether $\bar{q} > 0$ or $\bar{q} = 0$ (see the typical profiles in figures 1 and 2 respectively[2]). Therefore we write:

Remark 1 *In model (1) with δ given, for any true value $\bar{q} \geq 0$, the spectral likelihood almost surely has a local maximum at a $q > 0$.*

In general δ is unknown and needs to be estimated from the data. We will investigate this case next. Fernandez (1990) demonstrates that its spectral mle coincides with the moment estimator given by

$$\tilde{\delta} = T^{-1}(y_T - y_0), \tag{8}$$

[2] As $T \to \infty$ figure 1 will remain essentially unaffected while figure 2 will show profiles more and more compressed against the origin (dotted lines); hence if $\bar{q} = 0$, $\tilde{q} > 0$ for $T < \infty$, but $\tilde{q} \to 0$ as $T \to \infty$.

and that —after substituting δ by its mle (8)— the likelihood should now be written as (5) except that the summations run from $j=1$ —the zero frequency being dropped altogether since it involves a fixed term— and P_j represents the periodogram of the differences process $\{\Delta y_t\}$. The corresponding score function becomes proportional to

$$\mathcal{S}_2(\bar{q},q)=\frac{\beta_2(\bar{q},q)}{\alpha_2(\bar{q},q)}-\gamma_2(q), \tag{9}$$

$$\begin{aligned}
\alpha_2(\bar{q},q) &= \frac{1}{T}\sum_{j=1}^{\lfloor T/2\rfloor}\frac{\bar{q}+c_j+e_j}{q+c_j}\chi_j^2, \qquad \alpha_2(\bar{q},0)=\frac{1}{T}\sum_{j=1}^{\lfloor T/2\rfloor}[1+(\bar{q}+e_j)c_j^{-1}]\chi_j^2,\\
\beta_2(\bar{q},q) &= \frac{1}{T}\sum_{j=1}^{\lfloor T/2\rfloor}\frac{\bar{q}+c_j+e_j}{(q+c_j)^2}\chi_j^2, \qquad \beta_2(\bar{q},0)=\frac{1}{T}\sum_{j=1}^{\lfloor T/2\rfloor}c_j^{-1}[1+(\bar{q}+e_j)c_j^{-1}]\chi_j^2,\\
\gamma_2(q) &= \frac{1}{T-1}\sum_{j=1}^{T-1}\frac{1}{q+c_j}, \qquad \gamma_2(0)=\frac{1}{T-1}\sum_{j=1}^{T-1}c_j^{-1}=\frac{T+1}{12},
\end{aligned}$$

(see the typical profile in figure 3). Notice how the score does not shoot up when approaching zero from the right (as in figure 1) —instead it does so when nearing $-c_1=-4\sin^2(\pi/T)<0$— and hence $\arg(\mathcal{S}_2=0)$ is not neccesarily greater than zero. Then $\tilde{q}=\arg(\mathcal{S}_2=0)$ if $\arg(\mathcal{S}_2=0)>0$ but $\tilde{q}=0$ if $-c_1<\arg(\mathcal{S}_2=0)\le 0$.

Therefore, for a given true value $\bar{q}$, the probability of having a local maximum occurring at $q=0$ is

$$\begin{aligned}
p(\bar{q}) &= \Pr\left(\mathcal{S}_2(\bar{q},0)<0\right)\\
&= \Pr\left(\sum_{j=1}^{\lfloor T/2\rfloor}(c_j^{-1}-\frac{T+1}{12})(1+(\bar{q}+e_j)c_j^{-1})\chi_j^2<0\right).
\end{aligned} \tag{10}$$

For given T and $\bar{q}$, this probability can be computed by say Imhof (1961) routine or Davies (1980) algorithm (see table 1).

The special case when true q is precisely zero is of particular interest; actually we would also like to know the limiting probability of a local maximum occurring at $q=0$ when $\bar{q}=0$, so as to compare it to that obtained for time domain analogues. In this case, for finite T, table 1 shows that there will be a positive probability of finding a zero estimate regardless of whether $\bar{q}=0$ or not, but, on the other hand, it is easy to see that such probability will go to zero as $T\to\infty$ since $\sum_{j=1}^{\lfloor T/2\rfloor}(c_j^{-1}-\frac{T+1}{12})(1+e_jc_j^{-1})\chi_j^2$ is $O_p(T)$. Therefore we write:

Remark 2 *In model (1) with δ unknown, for any true value $\bar{q}\ge 0$, the spectral likelihood* may *have a local maximum at $q=0$, however the probability of such event vanishes as $T\to\infty$.*

Consequently, not only for small T the probabilities can be quite different, but also the *zero* limiting probability obtained in the frequency domain differs sharply from the 65.75% obtained in the time domain by Sargan & Bhargava (1983) for a unit root in an MA(1) model or Shephard & Harvey (1990) for the local-level model.

3 Sampling Distributions: Frequency vs. Time domain

Based on (10) and its parallel in the time domain (as in Shephard & Harvey 1990), table 1 compares the computed probabilities of obtaining a zero signal/noise estimate[3] in the frequency domain and in the time domain for given different true values and sample sizes (see also figure 4.) It clearly appears that obtaining zero estimates of q is far less probable in the frequency domain[4]. This means that deterministic components (true q=0) are identified more often in the time domain but, on the other hand, testing in the frequency domain could be more powerful.

Exploiting these formulae further (by way of setting $\bar{q} = 0$ in (9) for different values of $q > 0$) the probability density functions of the signal/noise estimator when the true value is actually zero (both in the frequency domain and time domain), can be compared in figure 5. As expected, the time domain distribution concentrates its probability near the zero boundary. In contrast, it is notable that, consistency notwithstanding, a greater-than-zero mode appears in the frequency domain distribution for samples of moderate size onwards. In particular, figure 5a shows that for samples betweeen 70 and 200 observations the frequency domain exhibits a mode around $q = 0.007$, concentrating a high probability around this value. Table 2 gives quantiles of these distributions[5]

and thus it may be helpful in testing for an MA unit root or, conversely, for the presence of a deterministic linear trend in non-seasonal time series.

Finally, the rate of consistency of the spectral mle turns out to be much slower than T^2-consistency exhibited in the time domain. In fact it can be shown that when the true signal-to-noise ratio $\bar{q} = 0$ then there is a local maximum of the likelihood function within a distance of $O(T^{-1})$ from the true value:

Remark 3 *For any $\alpha > 0$ there exist T_0 and $\epsilon > 0$ such that*

$$\Pr\left(\mathcal{S}(0, T^{-1}\epsilon) \geq 0\right) < \alpha, \quad \textit{for all } T > T_0; \qquad (\mathcal{S} = \mathcal{S}_1 \textit{ or } \mathcal{S}_2.)$$

Proof (see appendix A).

The intuition about this result comes from the fact that in practice we are not in the ideal world of circularity —where T^2-consistency could be achieved as in the time domain because the

[3]Strictly speaking a *local* maximum at zero. Simulation results suggest this is a good approximation to the desired probability; see also Shephard & Harvey (1990).

[4]Note also that estimating δ should push time-domain probabilities closer to one.

[5]see footnote 3.

likelihood would be exact— and that, in truth, the spectral density ordinates assumed by the approximate (spectral) likelihood approach the 'idealï ones needed for the likelihood to be exact at a rate of $O(T^{-1})$ only. This latter term will dominate the overall convergence rate of the estimator so that it actually turns out to be T-consistent instead.

These results are interesting as they may explain some of the differences between estimators from approximate likelihoods constructed in the frequency domain using the circularity assumption and exact time domain mle's; especially they help to explain how signal-to-noise ratios estimated from the spectral likelihood appear to give away the presence of deterministic components —and hence co-integration if applied, as in Fernandez (1993), to time-series regression residuals— far less often than when estimated in the time domain.

Appendix

A Consistency rates

The line of argument followed in this proof is similar in style to that used by Sargan & Barghava (1983) or Harvey & Shephard (1988). From (9), the scaled score when $\bar{q} = 0$ and $q = \epsilon / T^{\psi}$ is

$$\left(\frac{2}{T} \sum_j \frac{T^{\psi}(c_j + e_j)}{T^{\psi} c_j + \epsilon} \chi_j^2 \right)^{-1} \left(\sum_j \frac{T^{\psi}(c_j + e_j)}{(T^{\psi} c_j + \epsilon)^2} \chi_j^2 \right) - \sum_j \frac{1}{T^{\psi} c_j + \epsilon},$$

where χ_j^2 represent mutually independent chi-squared variates of 2 degrees of freedom. Applying a weak law of large numbers clearly $\text{plim}_{T \to \infty} (T/2)^{-1} \sum T^{\psi}(c_j + e_j)(T^{\psi} c_j + \epsilon)^{-1} \chi_j^2 = 2$ so that the scaled score is asymptotically equivalent to

$$X(\psi; \epsilon) = \sum_j \frac{T^{\psi} c_j (\chi_j^2 - 2) + T^{\psi} e_j \chi_j^2 - 2\epsilon}{2(T^{\psi} c_j + \epsilon)^2}.$$

The cumulants of the random variable X can be deduced from the chi-squared distribution, thus

$$\mathrm{E}\left(X(\psi; \epsilon) \right) = \sum_j \frac{T^{\psi} e_j - \epsilon}{(T^{\psi} c_j + \epsilon)^2}, \qquad \mathrm{V}\left(X(\psi; \epsilon) \right) = \sum_j \frac{T^{2\psi} (c_j + e_j)^2}{(T^{\psi} c_j + \epsilon)^4}.$$

In non-circular processes $e_j = O(T^{-1}) \; \forall j$ and therefore, to keep $\mathrm{E}(X)$ negative, ψ is to be at most equal to 1; then

$$\mathrm{E}\left(X(1; \epsilon) \right) \approx \sum_j \frac{T e_j - \epsilon}{(h_j + \epsilon)^2}, \qquad \mathrm{V}\left(X(1; \epsilon) \right) \approx \sum_j \frac{(h_j + T e_j)^2}{(h_j + \epsilon)^4},$$

with $h_j = 4\pi^2 j^2/T$ defined in the interval $[0, \pi^2 T]$. For any $\epsilon > Te_j \; \forall j$, Tchebycheff's inequality yields (with the summation starting either at $j = 0 : \mathcal{S} = \mathcal{S}_1$ or at $j = 1 : \mathcal{S} = \mathcal{S}_2$)

$$\lim_{T\to\infty} \Pr\left(T^{-1}\mathcal{S}(0, T^{-1}\epsilon) \geq 0\right) < k(\epsilon)/m^2,$$

where $m > (\epsilon - Te_j) \quad \forall j$ and $k(\epsilon)$ is bounded. Hence, choosing m sufficiently large so that also $m \geq \sqrt{(\sup k)/\alpha}$, we may write

$$\lim_{T\to\infty} \Pr\left(T^{-1}\mathcal{S}(0, T^{-1}\epsilon) \geq 0\right) < \alpha; \qquad (\mathcal{S} = \mathcal{S}_1 \text{ or } \mathcal{S} = \mathcal{S}_2)$$

and proposition 3 follows.

References

DAVIES,R.B.(1980), Algorithm AS 155. The Distribution of a Linear Combination of χ^2 Random Variables, *Applied Statistics,* **29,** 323–333.

ENGLE,R.F. and GRANGER,C.W.J. (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica,* **55,** 251–276.

FERNANDEZ-MACHO,F.J. (1986), *Estimation and Testing of Multivariate Structural Time Series Models,* Unpublished PhD Thesis, London School of Economics.

— (1990), Estimation and Testing of a Multivariate Exponential Smoothing Model, *Journal of Time Series Analysis,* **11,** 89–105.

— and HARVEY,A.C. (1990), Seemingly Unrelated Time Series Equations and a Test for Homogeneity, *Journal of Business & Economic Statistics,* **8,** 71–82.

— (1993), Spectral Maximum Likelihood Estimation in Regressions with Random-Walk-cum-Noise Disturbances: Testing for Co-integration, *mimeo, Dept. of Stats and Math Sci., LSE.*

FULLER,W.A. (1976), *Introduction to Statistical Time Series,* New-York: Wiley.

HARVEY,A.C. and DURBIN,J. (1986), The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modeling, *Journal of the Royal Statistical Society,* **A, 149,** 187–227.

HARVEY,A.C. and PETERS,S. (1990), Estimation Procedures for Structural Time Series Models, *Journal of Forecasting,* **9,** 89–108.

HARVEY,A.C. and SHEPHARD,N.G. (1988), On the Probability of Estimating a Deterministic Component in the Local Level Model, *mimeo,* Dpt. of Stat. and Math Sciences, London School of Economics.

IMHOF,J.P. (1961), Computing the Distribution of Quadratic Forms in Normal Variables, *Biometrika,* **48,** 419–426.

NELSON,C.R. and PLOSSER,C.I. (1982), Trends and Random Walks in Macroeconomic Time Series, *Journal of Monetary Economics,* **10,** 139–162.

NERLOVE,M., GRETHER,D.M. and CARVALHO,J.L. (1979), *Analysis of Economic Time Series,* New-York: Academic Press.

SARGAN,J.D. and BHARGAVA,A. (1983), Maximum Likelihood Estimation of Regression Models with First Order Moving Average Errors when the Root Lies on the Unit Circle, *Econometrica,* **51,** 799–820.

SHEPHARD,N.G. and HARVEY,A.C. (1990), On the Probability of Estimating a Deterministic Component in the Local Level Model, *Journal of Time Series Analysis,* **11,** 339–347.

Table 1: **Probabilities of obtaining a local maximum at $q = 0$.**

1a.- Frequency Domain.

T	True Value of q: 0	0.01	0.1	1	10
10	0.4298	0.4256	0.3922	0.2485	0.1349
20	0.3619	0.3487	0.2636	0.0876	0.0309
30	0.2992	0.2775	0.1658	0.0313	0.0081
40	0.2507	0.2223	0.1039	0.0120	0.0025
50	0.2129	0.1795	0.0655	0.0049	0.0008
60	0.1832	0.1433	0.0418	0.0021	0.0003
70	0.1591	0.1199	0.0269	0.0009	0.0001
100	0.1095	0.0684	0.0076	0.0001	0.0000
200	0.0397	0.0130	0.0002	0.0000	0.0000

1b.- Time Domain$^{(*)}$.

T	True Value of q: 0	0.01	0.1	1	10
10	0.6374	0.6118	0.4676	0.2146	0.1160
20	0.6474	0.5594	0.2870	0.0749	0.0316
30	0.6507	0.4865	0.1755	0.0300	0.0112
40	0.6515	0.4137	0.1099	0.0133	0.0038
50	0.6534	0.3487	0.0704	0.0062	0.0023
60	0.6541	0.2932	0.0441	...	...
70	0.6536	0.2467	0.0307	...	...
100	0.6554	0.1487	0.0099	...	...
200	0.6564	0.0322	0.0002	...	...

$^{(*)}$ Table II of Shephard and Harvey (1990) recalculated and extended.

Table 2: **Quantiles of distribution of signal/noise estimator (True value = 0)**

2a.- Frequency Domain.

T	75%	90%	95%	99%
50	0.07587	0.1413	0.202	0.4152
60	0.06024	0.1069	0.1479	0.2803
70	0.04976	0.0854	0.1155	0.2073
80	0.04224	0.07034	0.09395	0.1621
90	0.03679	0.05986	0.07863	0.1318
100	0.03246	0.05198	0.06758	0.1103
150	0.02000	0.03055	0.03838	0.05828
200	0.01476	0.02116	0.02610	0.03805

2b.- Time Domain.

T	75%	90%	95%	99%
50	0.002327	0.01102	0.02355	0.09052
60	0.002000	0.007260	0.01595	0.05814
70	0.001000	0.005627	0.01150	0.04071
80	0.0007569	0.004231	0.008553	0.03001
90	0.0005098	0.003128	0.006751	0.02298
100	0.0002589	0.002395	0.005160	0.01808
150	0.0001103	0.001125	0.002298	0.007599
200	0.0001021	0.001077	0.001536	0.004132

Figure 1: **Typical Profiles: known $\delta; \bar{q} > 0$.**

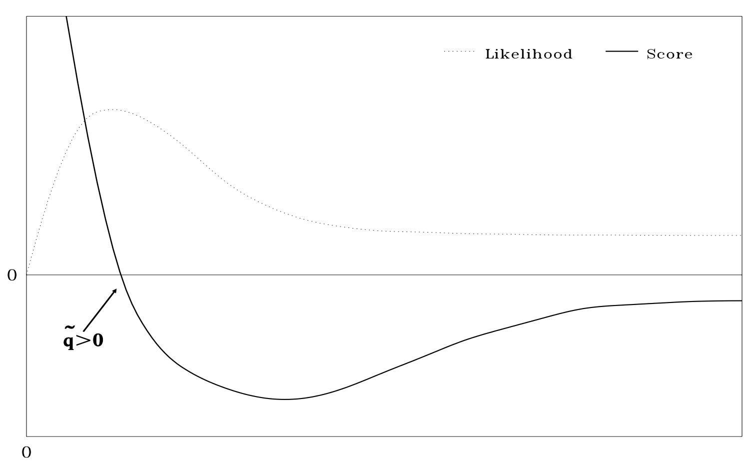


Figure 2: **Typical Profiles: known** $\delta; \bar{q} = 0$**.**

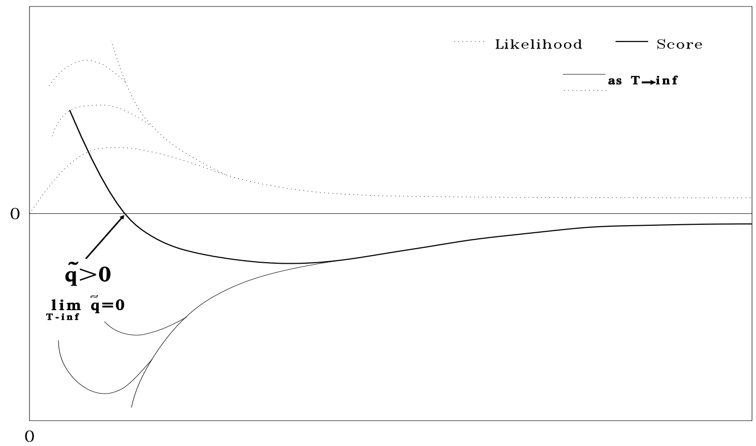


Figure 3: **Typical Profiles: unknown** $\delta; \bar{q} \geq 0$**.**

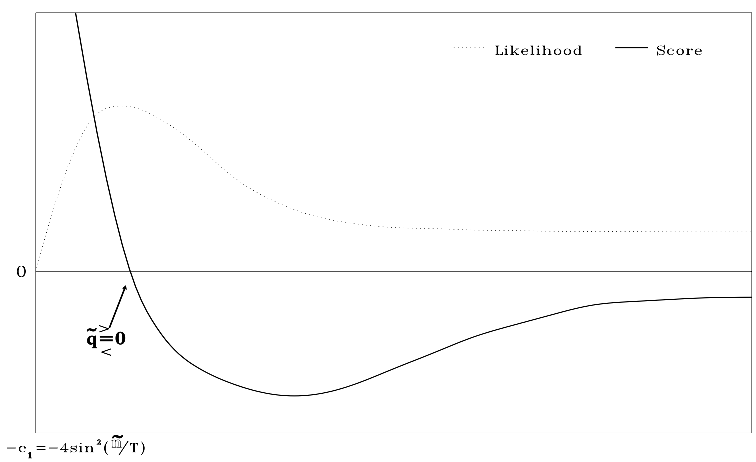
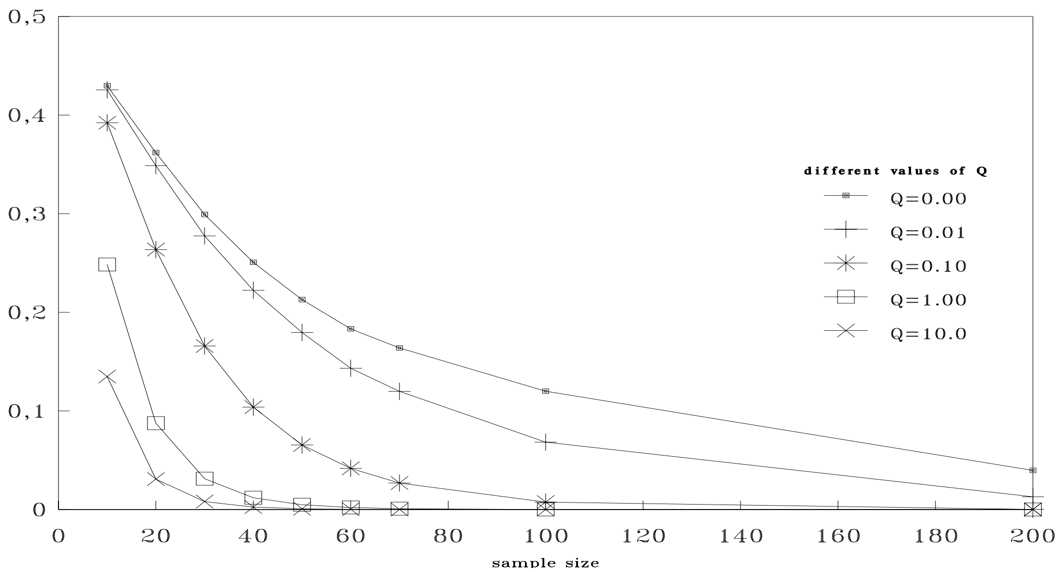


Figure 4: $\Pr(\tilde{q} = 0)$ **(against sample sizes).**

4a.- Frequency Domain.



4b.- Time Domain.

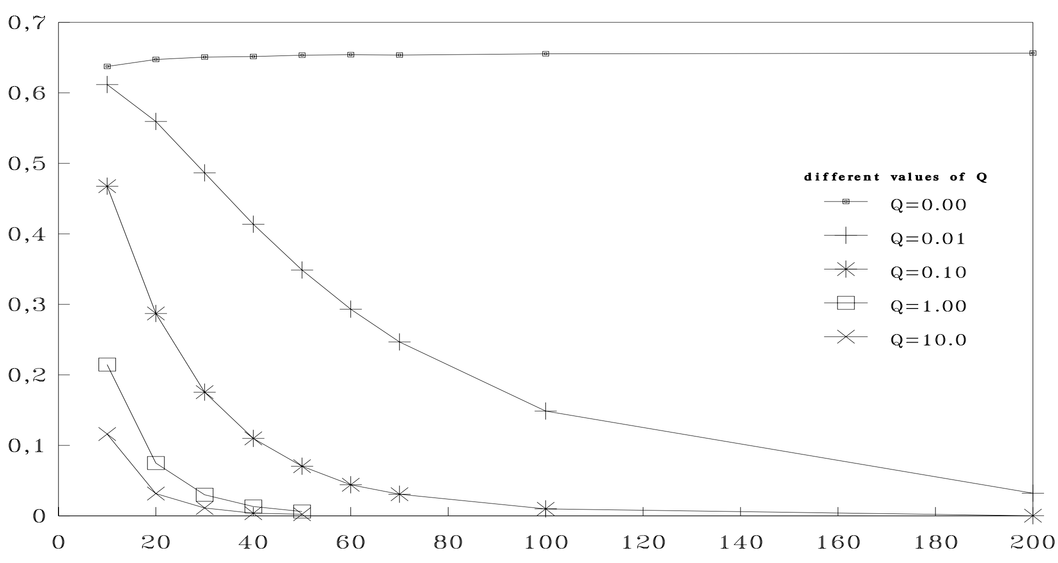
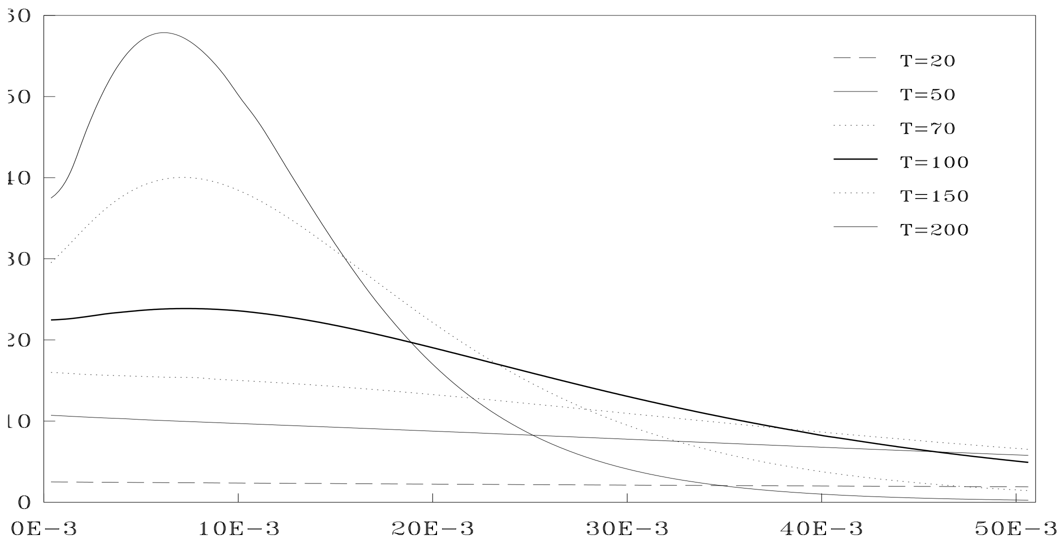


Figure 5: **Probability Density of $\tilde{q}/\bar{q} = 0$.**

5a.- (Non-circular) Frequency Domain.



5b.- Time Domain.