Size Distortion in the Analysis of Volatility and Covolatility Effects

Christian Gourieroux*

Joann Jasiak†

revised: February 6, 2009

(first version May 2007)

*CREST, CEPREMAP and University of Toronto , *e-mail*: gouriero@ensae.fr.

†York University, Canada, *e-mail*: jasiakj@yorku.ca

The authors gratefully acknowledge financial support of the Natural Sciences and Engineering Council of Canada (NSERC). Gourieroux thanks the Chair AXA/Risk Foundation: "Large Risk in Insurance".

Abstract

A Degeneracy in the Analysis of Volatility and Covolatility Effects

Let us assume that $\hat{A}_T$ is a consistent, asymptotically normal estimator of a matrix A (where T is the sample size), this paper shows that test statistics used in empirical work to test 1) the noninvertibility of A, i.e. $det A = 0$, 2) the positivite semi-definiteness $A >> 0$, have a different asymptotic distribution in the case where $A = 0$ than in the case where $A \neq 0$. Moreover, the paper shows that an estimator of A constrained by symmetry or reduced rank has a different asymptotic distribution when $A = 0$ than when $A \neq 0$. The implication is that inference procedures that use critical values equal to appropriate quantiles from the distribution when $A \neq 0$ may be size distorted. The paper points out how the above statistical problems arise in standard models in Finance in the analysis of risk effects.A Monte Carlo study explores how the asymptotic results are reflected in finite sample.

Keywords: Multivariate Volatility, Risk Premium, BEKK Model, Volatility Transmission, Identifiability, Boundary, Invertibility Test.
JEL number: C10, C32, G10, G12.

1 Introduction

In financial models, the risk on a set of assets is commonly represented by a volatility-covolatility matrix, while the risk effect on expected returns and future volatilities is often specified as an affine function of current and lagged realized volatilities and covolatilities. Under some specific hypotheses, the regularity conditions may not hold in this particular framework, and, as a consequence, the limiting distributions of some commonly used estimators and test statistics may differ from the standard ones.

There are two strands of literature that are directly concerned: the literature on risk premium, and the literature on multivariate ARCH models.

Let us first consider a simple risk premium model with 2 assets, called asset 1 and asset 2, and the following volatility matrix

$$\Sigma_t = \begin{pmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{22,t} \end{pmatrix}.$$

The expected return on asset 1 can be written as:

$$E_t(r_{1,t+1}) = r_{f,t+1} + a_1\sigma_{11,t} + 2b_1\sigma_{12,t} + c_1\sigma_{22,t} + a_1^*\sigma_{11,t-1} + 2b_1^*\sigma_{12,t-1} + c_1^*\sigma_{22,t-1},$$

where $r_{f,t}$ is the riskfree return, and coefficients (a_1, b_1, c_1), (a_1^*, b_1^*, c_1^*) are the elements of matrices $A = \begin{pmatrix} a_1 & b_1 \\ b_1 & c_1 \end{pmatrix}$ and A^*, respectively. This model allows us to estimate the ex-ante equity risk premium and to test the statistical significance and positivity of the risk premium. Technically, these two tests concern the significance and sign of matrix A (resp. A^*). Regarding the sign, there exists evidence that suggests that risk premium can be either positive or negative. In particular, Boudoukh et al (1993), Ostdiek (1985), Arnott, Ryan (2001), Arnott, Bernstein (2002), Chen, Guo, Zhang (2006), Walsh (2006) tested the positivity of the conditional risk premium using the method of instrumental variables and showed that risk premium can be of either sign, depending on the environment. The rank of risk premium is also unclear. The theory underlying the CAPM model suggests the existence of a relationship between the expected return and the variance of a single market portfolio that captures the entire effect of variances and covariances of all assets. This would imply that matrix A, in the above risk premium model, is not of full rank.

A similar ambiguity concerning the sign and rank of risk premium arises in foreign exchange markets [see e.g. Domowitz, Hakkio (1985), Macklem (1991) , Hakkio, Sibert (1995)]. The literature suggests that the sign of the foreign exchange real risk premium can vary depending on the ratio of market volatilities in both countries. The significance of risk premium is of economic interest too, as it has an important interpretation in the context of exchange rates. In particular, if the risk premium is zero, the forward exchange rate becomes an unbiased predictor of the future spot exchange rate.

In multivariate ARCH models, the expected future volatility is defined by linear functions of volatility-covolatility (see, e.g. Engle, Granger, Kraft (1984), Bollerslev, Engle, Wooldridge (1988), Bollerslev, Chou, Kroner (1992)). For example, the so-called vech-representation is:

$$V_t(r_{1,t+1}) = d + a_1\tilde{\sigma}_{11,t} + 2b_1\tilde{\sigma}_{12,t} + c_1\tilde{\sigma}_{22,t} + a_1^*\tilde{\sigma}_{11,t-1} + 2b_1^*\tilde{\sigma}_{12,t-1} + c_1^*\tilde{\sigma}_{22,t-1},$$

where $\tilde{\sigma}_{ij,t} = r_{i,t}r_{j,t},\ i,j = 1,2$, and as before, the remaining coefficients are elements of matrices A and A^*.

In this model, it is interesting to test the significance of lagged realized volatility, and the existence of a factor representation of realized volatility, as in the BEKK model (Baba, Engle, Kraft, Kroner (1990)). These tests directly concern the rank and sign of matrix A (resp. A^*), as it was the case in the risk premium model discussed in the previous paragraphs.

In this paper, we assume that matrix A is estimated from a sample of asset returns of size T, and that estimator $\hat{A}_T$ is a consistent, asymptotically normal estimator of A. Our study is focused on the tests of various hypotheses concerning matrix A, mainly for A of dimension 2×2, for clarity of exposition.

The hypotheses of interest discussed in this paper are:

1) the hypothesis of noninvertibility of matrix A;

2) the hypothesis that matrix A is positive semi-definite.

This last hypothesis is equivalent to the hypothesis of nonnegativity of the linear form $Tr(A\Sigma)$ (see Appendix 1, Lemma 1). Indeed, the linear form in volatilities-covolatilities that appears in the risk premium model and the vech-representation above can be rewritten as:

$$a\sigma_{11} + 2b_1\sigma_{12} + c\sigma_{22} = Tr\left[\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}\right] = Tr(A\Sigma),\ \text{say},$$

where Tr is the trace operator.

Moreover, we will also investigate

3) constrained estimation of A

when the constraint implies that the rank of A is less or equal to 1.

At a first sight, the above hypotheses tests and constrained estimation [1] seem quite standard. Indeed, the invertibility of matrix A (hypothesis 1) is usually tested from the singular value decomposition of the (asymptotically) Gaussian random matrix $\hat{A}_T$ [see Anderson (1989), Gourieroux, Monfort, Renault (1995), Bilodeau, Brenner (1999)]. As for hypothesis 2), the tests of matrix positivity are based on asymptotic tests of the following inequality restrictions $ac - b^2 \geq 0$, $a \geq 0$ [see e.g. Gourieroux, Monfort (1989), Wolak (1991)]. Finally, estimation of A under the hypothesis of reduced rank is commonly performed by a quasi-maximum likelihood method, as in the BEKK model [Engle, Kroner (1995), Jeantheau (1998), Comte, Lieberman (2000)].

The purpose of this paper is to point out the identifiability problems, boundary problems and degeneracies that may be encountered while performing the aforementioned tests and estimation in the framework of risk premium and multivariate ARCH models. In the presence of these effects, the asymptotic distributions of estimators and test statistics can be non-standard. This can render the outcomes of standard inference misleading and the stylized facts questionable. The degeneracies discussed in the paper concern some commonly used estimators and test statistics for which the true asymptotic distributions are derived. In particular, the asymptotic admissibility of the test statistics and their potential improvements are out of the scope of the present paper.

The paper is organized as follows. Section 2 considers the Wald test of non-invertibility of matrix A based on the estimated determinant det $\hat{A}_T$. It shows that in the degenerate case A=0 the Wald test statistic has a non-Gaussian distribution and that this distribution depends on the asymptotic variance of random matrix $\hat{A}_T$. Section 3 discusses the constrained estimation of A when it is not of full rank. We point out that when A=0, the distribution of the constrained estimator is non-standard. Section 4 considers the test of positive semi-definiteness, that is, of the hypothesis defined by inequality constraints $a \geq 0, c \geq 0, ac - b^2 \geq 0$. We show that when A=0, the standard asymptotic theory is no longer valid. The necessary adjustments are given for an unconstrained A [respectively, for A of reduced rank] under the maintained hypothesis. Finite sample properties of the standard test statistics in the degenerate case are presented in Section 5. Section 6 concludes.

[1] Similar problems arise in the so-called vech-diagonal multivariate ARCH models, such as $\sigma_{ij,t} = d_{ij} + a_{ij}\tilde{\sigma}_{ij,t}$, i,j=1,2, $i \leq j$. It is easy to check that the expected volatility-covolatility matrix is positive semi-definite, if and only if, the matrix $A = (a_{ij})$ is positive semi-definite. This condition is sufficient only for matrices of larger dimension (Silberberg, Pafka (2001)).

2 Invertibility Tests

Let us consider the test of invertibility of matrix A, based on the significance of its determinant. The null hypothesis is:

$$H_0 : \ (\det A = 0). \tag{2.1}$$

2.1 The unconstrained model

Let us consider a n $\times$ n matrix of parameters A, and its consistent, asymptotically Gaussian estimator $\hat{A}_T$. vec A denotes a vector of length n^2 obtained by stacking the columns of matrix A. We assume that:

$$\sqrt{T}[vec(\hat{A}_T) - vec(A)] \xrightarrow{d} N(0, \Omega), \tag{2.2}$$

where Ω is a $(n^2 \times n^2)$ invertible matrix and $\xrightarrow{d}$ denotes the convergence in distribution. $\hat{A}_T$ conveys all relevant information about A contained in the data. From now on, model (2.2) is referred to as the unconstrained (asymptotic) model.

2.2 Wald test statistic

A standard method for testing the null hypothesis H_0 in (2.1) is based on the estimated determinant $det\hat{A}_T$ and its asymptotic distribution obtained by applying the δ-method.

Since $\frac{\partial(\det A)}{\partial(vecA)} = vec[cof(A)]$, where $cof(A)$ is the (n $\times n$) matrix whose elements are the cofactors[2] of elements of A, we get:

$$\sqrt{T}(det\hat{A}_T - detA) \xrightarrow{d} N(0, vec[cof(A)]'\Omega\, vec[cof(A)]). \tag{2.3}$$

The Wald test statistic for testing the null hypothesis (2.1) is:

$$\hat{\xi}_T = \frac{\sqrt{T} det\hat{A}_T}{[vec[cof(\hat{A}_T)]'\hat{\Omega}_T\, vec[cof(\hat{A}_T)]]^{1/2}}, \tag{2.4}$$

where $\hat{\Omega}_T$ is a consistent estimator of Ω. If $vec[cof(A)] \neq 0$, this Wald statistic follows asymptotically a standard normal distribution and a critical region of the type $\{|\hat{\xi}_T| > 1.96\}$ defines a test at asymptotic level 5%.

[2]A cofactor is a determinant obtained by deleting the row and column of a given element of a matrix preceded by a + or - sign depending whether the element is in a + or - position.

2.3 The degenerate case

The standard asymptotic properties of the test are valid as long as $vec[cof(A)] \neq 0$, that is, if $A \neq 0$. Otherwise, the asymptotic properties of the Wald test statistic are significantly altered.

i) Asymptotic Properties of the Estimated Determinant

When $A = 0$, we have $\sqrt{T} vec(\hat{A}_T) \xrightarrow{d} vec(A_\infty) \sim N(0, \Omega)$, say. Thus we have: $det(\sqrt{T}\hat{A}_T) \xrightarrow{d} det(A_\infty)$, or equivalently

$$T^{n/2} det \hat{A}_T \xrightarrow{d} det(A_\infty). \tag{2.5}$$

When $n \geq 2$, the asymptotic behavior differs from the standard behavior, since:

i) the speed of convergence is $1/(T^{n/2})$ instead of $1/\sqrt{T}$, that is greater;

ii) the limiting distribution is not Gaussian, but instead, it is a determinant transformation of a multivariate Gaussian distribution.

ii) Asymptotic Properties of the Wald Test Statistic

Similarly, we can examine the test statistic $\hat{\xi}_T$ when A=0. Since $cof(\sqrt{T}\hat{A}_T) = T^{(n-1)/2} cof(\hat{A}_T)$, we see that $\hat{\xi}_T \xrightarrow{d} \xi(A_\infty)$, where

$$\xi(A_\infty) = \frac{det(A_\infty)}{\{vec[cof(A_\infty)]' \Omega \, vec[cof(A_\infty)]\}^{1/2}}. \tag{2.6}$$

iii) Comparison with the Literature

The degenerate case considered here does not belong to those discussed in Andrews (2001), in which some parameters are not identifiable under the null. In our framework, matrix A is always identifiable. This explains why the asymptotic distribution of the Wald statistic differs from the distribution derived by Andrews (2001).

This degeneracy cannot be disregarded or circumvented, for example, by introducing a sequence of null hypotheses indexed by the number T of observations, such as $H_{0,T} : [detA = 0, ||A|| > h(T)]$, where $||A||^2$ denotes the largest eigenvalue of AA' and $h(T)$ is strictly positive and tends to zero at an appropriate rate, when T tends to infinity [3]. Indeed, the hypothesis $H_0 : \{A = 0\}$ does not belong in the union of this sequence of hypothesis $H_{0,T}$, and hypothesis H_0 has often structural interpretations whereas the sequence $H_{0,T}$ does not. For instance, the test of H_0 allows for determining the autoregressive order of a multivariate ARCH model [4]. In the application to

[3] Such a methodology is followed in the test of switching regimes, for the parameter representing the unknown switching date [Andrews (1993)].

[4] See Andrews (2001), Francq, Zakoian (2006) for tests concerning the orders of univariate GARCH processes.

risk premium, the condition A=0 characterizes the hypothesis of nonpredictability of asset returns that is of economic interest.

2.4 Critical values

1) Asymptotic Size of the Test

The multiplicity of limiting distributions of the Wald test statistic under the null hypothesis suggests that a detailed analysis of the type I error is needed, as the condition of asymptotic similarity on the boundary condition is violated [see Hansen (2003)]. For instance, suppose that the null hypothesis is rejected when the Wald statistic $\hat{\xi}_T$ is larger in absolute value than the critical value c, and let us denote $\mathcal{A}_\prime$ the set of noninvertible matrices A.

The size of the test for a finite sample of length T is equal to:

$$\alpha_T(c) = sup_{A \in \mathcal{A}_\prime} P_A(\hat{\xi}_T > c),$$

and is reached for matrix A_T^* in $\mathcal{A}_\prime$.

Then, it is possible to define

(*) the asymptotic null rejection probability as:

$$\alpha_\infty(c) = sup_{A \in \mathcal{A}_\prime} \lim_{T \to \infty} P_A(\hat{\xi}_T > c);$$

(**) the asymptotic size of the test as:

$$\tilde{\alpha}_\infty(c) = \lim_{T \to \infty} \alpha_T(c) = \lim_{T \to \infty} sup_{A \in \mathcal{A}_\prime} P_A(\hat{\xi}_T > c).$$

In the sequel, we assume that $\lim_{T \to \infty}$ and $sup_{A \in \mathcal{A}_\prime}$ can commute, which implies a "uniform convergence" condition of the finite sample distribution of $\hat{A}_T$ towards its asymptotic Gaussian distribution.

Assumption A.1:

The asymptotic size of the test is equal to the asymptotic null rejection probability.

In the applications, Assumption A.1. has to be verified case by case according to the type of asymptotically Gaussian estimator $\hat{A}_T$ which is used.

Under Assumption A.1, the asymptotic size of the test is:

$$\alpha_\infty(c) \quad = \quad sup_{H_0} \lim_{T \to \infty} P_A[|\hat{\xi}_T| > c]$$

$$= \quad sup[sup_{A:detA=0,A\neq 0} \lim_{T\to\infty} P_A(|\hat{\xi}_T| > c), sup_{A=0} \lim_{T\to\infty} P_A(|\hat{\xi}_T| > c)]$$
$$= \quad sup[(P(|X| > c), P(|\xi(A_\infty)| > c)] \text{ (where } X \sim N(0,1)).$$

By inverting this relationship, we deduce the critical value for an asymptotic size $\alpha_\infty = \alpha$:

$$c(\alpha) = Max[\Phi^{-1}(1-\alpha/2), Q(\alpha, \Omega)],$$

where Φ is the cdf of the standard normal, and $Q(\alpha, \Omega)$ is the quantile computed from:

$$P[|\xi(A_\infty)| > Q(\alpha, \Omega)] = \alpha, \tag{2.7}$$

where $vec(A_\infty) \sim N(0, \Omega)$.

Function Q is too complicated to be calculated analytically, but the value $Q(\alpha, \Omega)$ can be easily approximated by Monte-Carlo simulations. Let us denote by $\hat{\Omega}_{0T}$ an estimator of Ω, which is consistent under the null and by $\hat{Q}(\alpha, \hat{\Omega}_{0T})$ the associated value of Q derived by simulations. The critical value will be chosen as $\hat{c}(\alpha) = Max[\Phi^{-1}(1-\alpha/2), \hat{Q}(\alpha, \hat{\Omega}_{0T})]$.

ii) Comparison with Sequential Procedures

Under Assumption A.1, the procedure above provides the correct asymptotic size of the test of the null hypothesis $H_0 : \{\det A = 0\}$. It is an alternative to the sequential procedures described below, which are asymptotically size distorted.

i) A two-step procedure can be as follows. In the first step, we consider a Fisher statistic F for testing the hypothesis $H_0^* : \{A = 0\}$ with critical value f_{α_0}, say, corresponding to level α_0. If $F < f_{\alpha_0}$, the null hypothesis H_0 is accepted. Otherwise, in the second step we perform a test based on the determinant at level α_1, and accept H_0, if $\xi_T < \Phi^{-1}(1-\alpha_1/2)$. The critical region of the sequential test is:

$$W = \{F > f_{\alpha_0}, \xi_T > \Phi^{-1}(1-\alpha_1/2)\}.$$

For a given choice of α_0, α_1, the asymptotic size[5] of this test is equal to

$$Sup_{A:det(A)=0} \lim_{T\to\infty} P[W > f_{\alpha_0}, \xi_T > \Phi^{-1}(1-\alpha_1/2)].$$

The asymptotic size can be bounded by a known function of α_0, α_1, but depends on α_0, α_1 and Ω, in general. Thus, this sequential test can be asymptotically size distorted.

[5]Assumed equal to the asymptotic null rejection probability.

ii) Another sequential procedure can be based on the analysis of the rank of matrix A [see e.g. Anderson (1989), Gill, Lewbel (1992), Cragg, Donald (1996), (1997), Bilodeau, Bremer (1999), Robin, Smith (2000)]. Indeed, for a matrix of dimension $(n \times n)$, we know that

$$\begin{aligned} H_0 : \{detA = 0\} &= \{rank(A) = 0\} \cup \{rank(A) = 1\} \cup \ldots \{rank(A) = n-1\} \\ &= \{rank(A) < n\}. \end{aligned}$$

Thus, we can first test if $rankA = 0$; then, if this hypothesis is rejected, we test if $rankA = 1$, etc. As above, the asymptotic size of this sequential test can be easily bounded, but its exact value is difficult to derive. The interpretation in terms of rank shows that a) a reason for the degenerate asymptotic behavior of statistic $\hat{\xi}_T$ is that the null hypothesis H_0 is a union of elementary null hypotheses $\{rank(A) = p\}$; b) the only elementary hypothesis that causes the degeneracy is $\{rank(A) = 0\}$, whereas the other elementary hypotheses $\{rank(A) = p\}$, $p = 1, ..., n-1$ have been jointly accommodated in the single statistic $\hat{\xi}_T$.

2.5 Symmetric matrix A of dimension (2,2)

Let us consider the estimator of a square (2,2) symmetric matrix $\hat{A}_T = \begin{pmatrix} \hat{a}_T & \hat{b}_T \\ \hat{b}_T & \hat{c}_T \end{pmatrix}$ and its convergence limit $A_\infty = \begin{pmatrix} a_\infty & b_\infty \\ b_\infty & c_\infty \end{pmatrix}$. The aim of this section is to derive the asymptotic critical values of the Wald test for any possible matrix Ω. Matrix Ω contains 6 different elements. First, we show that the critical values depend on Ω by only three parameters and the sign. This finding allow us to simplify the display of critical values.

The test statistic $\xi(A_\infty)$ is such that $\xi(PA_\infty P') = \xi(A_\infty)$, for any matrix P of dimension $(n \times n)$ (see Proposition A.1, ii) in Appendix 2). We infer that the quantiles $Q(\alpha, \Omega)$ and $Q(\alpha, \Omega(P))$ are identical if $\Omega = V[vec(A_\infty)]$ and $\Omega(P) = V[vec(PA_\infty P')]$, for any P. By choosing an appropriate linear transformation P, we show in Appendix 2, b) that the quantiles $Q(\alpha, \Omega)$, $\forall \Omega$, depend in fact, on a number of parameters much smaller than the number of elements in Ω.

Proposition 1: Up to a transformation $A_\infty \longrightarrow PA_\infty P'$, matrix Ω can be defined as:

$$\Omega = Var \begin{pmatrix} a_\infty \\ b_\infty \\ c_\infty \end{pmatrix} = \begin{pmatrix} 1 & 0 & \epsilon\rho^2 \\ 0 & \gamma^2 & 0 \\ \epsilon\rho^2 & 0 & 1 \end{pmatrix},$$

where parameters ρ and γ are nonnegative, $\rho < 1$, and ϵ is equal to +1 or -1, according to the sign of correlation between a_∞ and c_∞.

Thus, the set of admissible quantiles $\{Q(\alpha,\Omega),\Omega>>0\}$, where $\Omega>>0$ means that the matrix is symmetric, positive semi-definite, coincides with the set of quantiles $\{Q[\alpha,\Omega(\epsilon,\rho,\gamma)],\epsilon=\pm1,0<\rho<1,\gamma>0\}$.

The Wald test statistic with $\Omega(\epsilon,\rho,\gamma)$ is:

$$\begin{aligned}\xi(A_\infty) &= \frac{a_\infty c_\infty - b_\infty^2}{\sqrt{(c_\infty,-2b_\infty,a_\infty)\Omega(\epsilon,\rho,\gamma)(c_\infty,-2b_\infty,a_\infty)'}}\\ &= \frac{a_\infty c_\infty - b_\infty^2}{\sqrt{c_\infty^2+a_\infty^2+2\epsilon\rho^2 c_\infty a_\infty+4b_\infty^2\gamma^2}}. \qquad (2.8)\end{aligned}$$

Table 1 provides the upper quantiles at 10%, 5% and 1% of the variable $|\xi(A_\infty)|$ for different values of parameters ρ,γ and $\epsilon=+/-1$. The quantiles are obtained from Monte-Carlo experiments with 5000 replications. They can be directly compared to the critical values 1.64, 1.96, 2.57 of the standard normal distribution, which correspond to the case when $detA=0$ with $A\neq0$. We observe that all these values are smaller than their Gaussian counterparts. This implies that, under Assumption A.1. for a (2,2) symmetric matrix A the asymptotic size of the standard Wald test does not need to be corrected for degeneracy A=0, but likely, the magnitude of size distortion depends on the dimension of matrix A. Moreover, we show that such a size correction is needed for other inference on matrix A.

3 Constrained Estimation of A

3.1 The Example of BEKK model

To ensure the positivity of volatility matrix $H_t=V_t(r_{t+1})$, the multivariate GARCH literature (Engle, Kroner (1995)) proposed the following constrained specification [6]:

$$H_t = C_0 + \sum_{j=1}^{p} M_j H_{t-j} M_j' + \sum_{k=1}^{q} N_k r_{t-k} r_{t-k}' N_k', \text{ say,}$$

where M_j, N_k, C_0 are (n,n) matrices and $C_0>>0$. Accordingly, the volatility of asset i is:

$$h_{iit} = c_{0,ii} + \sum_{j=1}^{p} M_{ij} H_{t-j} M_{ji}' + \sum_{k=1}^{q} N_{ik} r_{t-k} r_{t-k}' N_{ki}',$$

where M_{ij} (resp. N_{ik}) is the i^{th} row of M_j (resp. N_k). A component of the first sum on the right-hand side is of the form:

[6] For ease of exposition, we introduced only 1 positive component per lag.

$$M_i H M_i' = Tr(M_i H M_i') = Tr(M_i' M_i H) = Tr(A_i H), \text{ say,}$$

where $A_i = M_i' M_i$ is of rank less or equal to 1.

Under a BEKK specification, the estimation of matrix A_i has to be performed under constraints. A common approach consists in optimizing a quasi-likelihood function with respect to parameters M (and N) [see e.g. Engle, Kroner (1995), Comte, Lieberman (2003), Iglesias, Phillips (2005)]. Let us consider matrix A of dimension two [7], $A = \begin{pmatrix} m_1^2 & m_1 m_2 \\ m_1 m_2 & m_2^2 \end{pmatrix}$. Due to a lack of identifiability of parameter M, the following two difficulties arise:

i) First, there is a problem of global identifiability since the same matrix A is obtained for M and -M. To solve this problem, it is common to use the following change of parameters:

$$A = m_1^2 \begin{pmatrix} 1 & m_2/m_1 \\ m_2/m_1 & (m_2/m_1)^2 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ \beta \end{pmatrix} (1, \beta), \text{ say,} \tag{3.1}$$

where $\alpha = m_1^2 \geq 0, \beta = m_2/m_1$ (whenever $m_1 \neq 0$, or equivalently $\alpha \neq 0$).

ii) Second, there is a problem of local identifiability at $A = 0$. The reason is that the Jacobian

$$\frac{\partial \, vechA}{\partial (m_1, m_2)'} = \begin{pmatrix} 2m_1 & 0 \\ m_2 & m_1 \\ 0 & 2m_2 \end{pmatrix}$$

is of rank 2, except when $A = 0$.

The asymptotic theory established for multivariate BEKK models doesn't hold for the estimators of parameters α and β defined in (3.1), because it assumes the identifiability of parameter M [see Assumption A.4 in Comte, Lieberman (2003)]. To overcome this difficulty Engle, Kroner (1995) (Proposition 2.1) introduce the identifiability condition $m_1 > 0$. This condition eliminates both the global and local identifiability problems.

In the next section, we derive the true asymptotic distributions of the minimum distance estimators of α and β based on a consistent, and asymptotically normal estimator of A. For the application to BEKK model, we assume that the unconstrained quasi-maximum likelihood estimator of A is asymptotically normal. This, in turn, requires some additional assumptions on the BEKK model, such as the presence of at least one non-zero ARCH effect [$N_{ki} \neq 0$ for at least one index k] to avoid another degeneracy pointed out in Andrews (2001).

[7] The results can be easily extended to matrix A of dimension $(n \times n)$ and of rank less or equal to 1.

3.2 The constrained estimator

Let us now assume that matrix A is symmetric and of reduced rank. Then we can write $A = \alpha \begin{pmatrix} 1 \\ \beta \end{pmatrix} (1, \beta)$, where α and β are unconstrained [8].

The constrained estimator of A based on $\hat{A}_T$ is the solution of the following minimization:

$$(\hat{\alpha}_T, \hat{\beta}_T) = \arg min_{\alpha,\beta} (\hat{a}_T - \alpha, \hat{b}_T - \alpha\beta, \hat{c}_T - \alpha\beta^2) \tilde{\Omega}_T^{-1} \begin{pmatrix} \hat{a}_T - \alpha \\ \hat{b}_T - \alpha\beta \\ \hat{c}_T - \alpha\beta^2 \end{pmatrix}. \quad (3.2)$$

The objective function (3.2) is defined for all values of parameters α, β. However, the stochastic coefficients involved in the objective function cannot be normalized uniformly with respect to the true matrix A. Therefore, Assumption 3 in Andrews (1999), p. 1349, is not satisfied and new asymptotic results need to be derived.

The objective function can be concentrated with respect to α. Then, the solution in α for a given β is:

$$\alpha(\beta) = < vech\hat{A}_T, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} > / < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >,$$

where $<,>$ denotes the inner product associated with $\tilde{\Omega}_T^{-1}$.

The concentrated objective function is:

$$\Psi_T(\beta) = < vech\hat{A}_T, vech\hat{A}_T > - \left[(< vech\hat{A}_T, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >) \right]^2 / < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} > . \quad (3.3)$$

The optimization of the concentrated objective function yields a finite solution (see Appendix 4).

Since the first-order condition is:

$$< vech\hat{A}_T, \begin{pmatrix} 0 \\ 1 \\ 2\beta \end{pmatrix} >< \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} > - < vech\hat{A}_T, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >< \begin{pmatrix} 0 \\ 1 \\ 2\beta \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >= 0, \quad (3.4)$$

the solution that minimizes (3.3) is a root of a polynomial of degree 5.

[8] We do not assume a priori that A is positive semi-definite.

3.3 Asymptotic distribution of the constrained estimator

When A is not equal to zero (i.e. if $\alpha \neq 0$), the standard asymptotic theory holds and we have:

$$\sqrt{T}\left[\begin{pmatrix} \hat{\alpha}_T \\ \hat{\beta}_T \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right] \xrightarrow{d} N[0, (J(\alpha,\beta)\tilde{\Omega}^{-1}J(\alpha,\beta)')^{-1}],$$

where the Jacobian matrix is $J(\alpha,\beta) = \begin{pmatrix} 1 & \beta & \beta^2 \\ 0 & \alpha & 2\alpha\beta \end{pmatrix}$.

When A=0, the Jacobian matrix is of rank 1, and the standard asymptotic theory is no longer valid. Let us now consider this case. It follows from (3.4) that $\hat{\beta}_T$ is a solution of

$$\begin{aligned} & Max_\beta \quad \left[< vech\hat{A}_T, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >\right]^2 / < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} > \\ \Longleftrightarrow \quad & Max_\beta \left[< vech(\sqrt{T}\hat{A}_T), \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >\right]^2 / < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} > . \end{aligned}$$

As a consequence, $\hat{\beta}_T$ tends to a limit β_∞, which is a solution to the optimization:

$$Max_\beta \left[< vech(A_\infty), \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >\right]^2 / < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} > . \tag{3.5}$$

Similarly, we note that:

$$\begin{aligned} \sqrt{T}\hat{\alpha}_T &= \sqrt{T}\alpha(\hat{\beta}_T) \\ &= < vech(\sqrt{T}\hat{A}_T), \begin{pmatrix} 1 \\ \hat{\beta}_T \\ \hat{\beta}_T^2 \end{pmatrix} > / < \begin{pmatrix} 1 \\ \hat{\beta}_T \\ \hat{\beta}_T^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \hat{\beta}_T \\ \hat{\beta}_T^2 \end{pmatrix} > \end{aligned}$$

tends to a limit

$$\alpha_\infty =< vech(A_\infty), \begin{pmatrix} 1 \\ \beta_\infty \\ \beta_\infty^2 \end{pmatrix} > / < \begin{pmatrix} 1 \\ \beta_\infty \\ \beta_\infty^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta_\infty \\ \beta_\infty^2 \end{pmatrix} > . \tag{3.6}$$

Proposition 3 summarizes the above discussion.

Proposition 3

If A=0, then $(\sqrt{T}\hat{\alpha}_T, \hat{\beta}_T) \xrightarrow{d} (\alpha_\infty, \beta_\infty)$, where $(\alpha_\infty, \beta_\infty)$ is a complicated nonlinear transformation of the Gaussian vector, derived from (3.5),(3.6).

Note that parameter β is not identifiable when A=0. Nevertheless its estimator $\hat{\beta}_T$ admits a limiting distribution.

The asymptotic limiting distributions of test statistics for α and β are non-standard too. For instance, the t-statistic for the test of significance of parameter α is:

$$\hat{\eta}_T^\alpha = \sqrt{T}\hat{\alpha}_T / \hat{\sigma}_{\alpha,T},$$

where $\hat{\sigma}_{\alpha,T}$ is the square root of the first diagonal element of the matrix $[J(\hat{\alpha}_T, \hat{\beta}_T)\tilde{\Omega}_T^{-1} J(\hat{\alpha}_T, \hat{\beta}_T)']^{-1}$. When $A \neq 0$, this statistic tends in distribution to a standard normal. When A=0, statistic $\hat{\eta}_T^\alpha$ tends to:

$$\eta_\infty^\alpha = < vech(A_\infty), \begin{pmatrix} 1 \\ \beta_\infty \\ \beta_\infty^2 \end{pmatrix} > / \sigma_{\alpha,\infty}, \tag{3.7}$$

where $\sigma_{\alpha,\infty}$ is the square root of the first diagonal element of the random matrix
$\Sigma_\infty = [J(\alpha_\infty, \beta_\infty)\tilde{\Omega}^{-1} J(\alpha_\infty, \beta_\infty)']^{-1}$.
Similarly, the t-statistic for the test of significance of parameter β,

$$\hat{\eta}_T^\beta = \sqrt{T}\hat{\beta}_T / \hat{\sigma}_{\beta,T}$$

tends to

$$\eta_\infty^\beta = \beta_\infty / \sigma_{\beta,\infty}, \tag{3.8}$$

where $\sigma_{\beta,\infty}$ is the square root of the second diagonal element of Σ_∞.
Table 2 presents the quantiles at 10%, 5%, 1% of the distribution of variables $|\eta_\infty^\alpha|$ and $|\eta_\infty^\beta|$, respectively, calculated for Gaussian matrices introduced in Section 2. The quantiles have been obtained by simulations with 5000 replications.

The quantiles associated with the t-statistic for α are less sensitive to parameters ρ and γ than the quantiles associated with the t-statistics for β. Both sets of quantiles are much more sensitive to parameter γ than to other parameters. Moreover, the quantiles differ significantly from the

Gaussian quantiles 1.64, 1.96, 2.57, especially for parameter β. In particular, the critical values exceed significantly the critical values from the standard normal distribution.

Figure 1 shows the distribution of β_∞ for $\rho = 0, \gamma = 1$. For $\rho = 0, \gamma = 1$, β_∞ is the solution of $Max_\beta (a_\infty + b_\infty \beta + c_\infty \beta^2)^2/(1+\beta^2+\beta^4)$, where $a_\infty, b_\infty, c_\infty$ are independent standard normal. Since

$$\begin{aligned} \beta_\infty(-a_\infty, -b_\infty, -c_\infty) &= \beta_\infty(a_\infty, b_\infty, c_\infty), \\ \beta_\infty(c_\infty, b_\infty, a_\infty) &= 1/\beta_\infty(a_\infty, b_\infty, c_\infty), \end{aligned}$$

the distribution of β_∞ is symmetric and invariant with respect to transformation $\beta_\infty \to 1/\beta_\infty$. This explains the shape of the distribution displayed in Figure 1, with a mode at 0 and very heavy tails.

[Insert Figure 1: Distribution of β_∞]

4 Positivity Test

Let us now focus on the test of positivity for a symmetric matrix A. This test depends on the maintained hypothesis, that is, on whether we assume "A unconstrained", or "A of reduced rank". Both cases are discussed below.

4.1 A unconstrained

A common approach to testing matrix positivity is as follows. The null hypothesis is written as $H_0 : \{a \geq 0, c \geq 0, ac - b^2 \geq 0\}$, and the test of these inequality restrictions is performed along the lines developped [9] by [Gourieroux, Holly, Monfort (1980), (1982), Kodde, Palm (1986), Gourieroux, Monfort (1989), Wolak (1991)]. However, in the presence of a degeneracy due to A=0, this standard technique cannot be applied. The reason is that it requires the Jacobian of the constraints, that is, $(a, b, c) \to (a, ac - b^2)$ to be of full rank on the boundaries of the null hypothesis. For A=0, however, the Jacobian $\begin{pmatrix} 1 & 0 & 0 \\ c & -2b & a \end{pmatrix}$ is of reduced rank.

Intuitively, the degeneracy can be explained as follows. The positivity condition involves three restrictions and the null hypothesis should be written as $H_0 : \{a \geq 0, c \geq 0, ac - b^2 \geq 0\}$. If either a (resp. c) is strictly positive, than condition $ac - b^2 \geq 0$ implies that c (resp. a) is nonnegative.

[9] see e.g. example iv) in Andrews, (1996), p. 705.

Thus, one of the two first inequalities seems to be redundant. In fact, this is not the case. For instance, the restrictions $a \geq 0, ac - b^2 \geq 0$ are satisfied for $A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$, which is not positive semi-definite.

Let us now consider the asymptotic properties of the likelihood ratio test. The log-likelihood function of the (asymptotic) unconstrained model is :

$$L_T(A) = T[-\log 2\pi - \frac{1}{2}\log det\tilde{\Omega}_T - \frac{1}{2}vech(\hat{A}_T - A)'\tilde{\Omega}_T^{-1}vech(\hat{A}_T - A)], \tag{4.1}$$

where vech denotes the vec-half operator. The likelihood ratio statistic for testing the positivity hypothesis is:

$$\begin{aligned} \xi_T^P &= 2(Max_A L_T(A) - Max_{A:A>>0} L_T(A)) \\ &= Min_{A:A>>0} T vech(\hat{A}_T - A)'\tilde{\Omega}_T^{-1}vech(\hat{A}_T - A). \end{aligned} \tag{4.2}$$

The estimator of matrix A constrained by the positivity condition can be equal to either of the three following:

i) $\hat{A}_T$, when $\hat{A}_T >> 0$;

a solution to (4.2), which can be either:

ii) a positive semi-definite matrix of rank 1;

iii) 0.

Under standard regularity conditions, the maximum value of type I error under the null is attained for A=0, and is computed from a weighted mixture of chi-square distributions, with weights equal to the probabilities of the three outcomes i), ii), iii) evaluated under A=0.

For A=0 however, some identification problems arise, as shown in the previous sections. Let us consider the asymptotic behavior of the likelihood ratio statistic when A=0. Since the set of positive semi-definite matrices is a positive cone, we get:

$$\begin{aligned} \xi_T^P &= Min_{A:A>>0} T vech(\hat{A}_T - A)'\tilde{\Omega}_T^{-1}vech(\hat{A}_T - A) \\ &= Min_{A:A>>0} vech(\sqrt{T}\hat{A}_T - A)'\tilde{\Omega}_T^{-1}vech(\sqrt{T}\hat{A}_T - A) \\ \xrightarrow{d} \xi_\infty^P &= Min_{A:A>>0} vech(A_\infty - A)'\tilde{\Omega}^{-1}vech(A_\infty - A). \end{aligned} \tag{4.3}$$

Thus, (4.3) defines an asymptotic optimization criterion under A=0. There are 3 regimes distinguished by the admissible values of that objective function:

Value in regime i) : $\xi_{\infty}^{1,P} = 0$;

Value in regime ii) : $\xi_{\infty}^{2,P} = vech(A_{\infty} - A_{\infty}^{0})'\tilde{\Omega}^{-1}vech(A_{\infty} - A_{\infty}^{0})$,

where $vech(A_{\infty}^{0})' = (\alpha_{\infty}, \alpha_{\infty}\beta_{\infty}, \alpha_{\infty}\beta_{\infty}^{2})$;

Value in regime iii) : $\xi_{\infty}^{3,P} = vech(A_{\infty})'\tilde{\Omega}^{-1}vech(A_{\infty})$.

The asymptotic probabilities of these regimes are denoted by $\pi_{\infty}^{1}, \pi_{\infty}^{2}, \pi_{\infty}^{3}$.

Let us now consider the type I error. We get

$$sup_{A>>0} \lim_{T\to\infty} P[\xi_T^P > c] = sup[sup_{A>>0,A\neq 0} \lim_{T\to\infty} P[\xi_T^P > c], P_{A=0}[\xi_{\infty}^P > c]].$$

By standard asymptotic theory underlying the tests of inequality constraints [see e.g. Gourieroux, Holly, Monfort (1980), Gourieroux, Monfort (1989), Wolak (1991)], the first component $sup_{A>>0,A\neq 0} \lim_{T\to\infty} P[\xi_T^P > c]$ is bounded from above by the survival function corresponding to a mixture of chi-square [10]:

$$\pi_{\infty}^{1}\chi^2(0) + \pi_{\infty}^{2}\chi^2(2) + \pi_{\infty}^{3}\chi^2(3).$$

This survival function has to be compared with the survival function of ξ_{∞}^{P} under A=0. This survival function is of the type:

$$\pi_{\infty}^{1}\chi^2(0) + \pi_{\infty}^{2}Q_{\infty} + \pi_{\infty}^{3}\chi^2(3),$$

where Q_{∞} denotes the asymptotic distribution of $\xi_{\infty}^{2,P}$. As in the previous sections, the limiting distribution Q_{∞} and the probabilities of the regimes can be easily obtained from simulations.

4.2 A of reduced rank

Section 3 considered the estimation of A when the rank of matrix A is less or equal to 1. In this parametric framework, the positivity hypothesis can be written as $H_0 : (\alpha \geq 0)$. It is usually tested by a one-sided test based on the t-statistic $\hat{\eta}_T^{\alpha}$. As shown in Section 3, the asymptotic distribution of this test statistic is standard normal, except when $\alpha = 0$ (that is A=0). We provide in Table 3 the one-sided critical value, that is the lower quantile of η_{∞}^{α} at 1%, 5%, 10%, derived by simulation with 5000 replications.

[10] Under regime ii), the standard theory implies a mixture of $\chi^2(1)$ and $\chi^2(2)$, which is bounded from above by a $\chi^2(2)$.

5 Finite Sample Properties

The previous sections were focused on the asymptotic distributions of test statistics. These distributions can be significantly different from the finite sample distributions.

To study the finite sample properties of standard test statistics, we generate three samples of iid standard Gaussian returns $(r_{1,t}, r_{2,t})'$, that are $IIN(0, Id)$, where Id denotes the identity matrix. The number of observations in each sample is T=50, 100, 200. Next, we consider the following regressions:

Regression 1: $r_{1,t} = d + ar_{1,t-1}^2 + 2br_{1,t-1}r_{2,t-1} + cr_{2,t-1}^2 + v_t$;

Regression 2: $r_{1,t}^2 = d + ar_{1,t-1}^2 + 2br_{1,t-1}r_{2,t-1} + cr_{2,t-1}^2 + v_t$.

The first regression is a model with a bivariate risk premium, while the second one is a volatility transmission model.

For each regression, we determine the finite sample distributions of $\hat{\xi}_T, \hat{\eta}_T^{\alpha}, \hat{\eta}_T^{\beta}$, where the Wald statistics are derived from the OLS estimators of a,b,c with the OLS estimated variance-covariance matrix $\tilde{\Omega}_T$. The distributions of $\hat{\xi}_T$ for the two regressions are displayed in Figures 2a-2b. We observe fat tails, and different limiting distributions for each of the two regressions due to the differences between the limiting OLS covariance matrices for the two regressions (see Section 2.5).

[Insert Figures 2a, 2b : Finite sample distribution of $\hat{\xi}_T$]

Let us now consider the finite sample distributions of the t-ratios for α and β. All the distributions feature fat tails due to the stochastic variance in the denominator of the t-ratio.

[Insert Figures 3a, 3b : Finite sample distribution of $\hat{\eta}_T^{\alpha}$]

[Insert Figures 4a, 4b : Finite sample distribution of $\hat{\eta}_T^{\beta}$]

6 Concluding Remarks

The paper derives the limiting distributions of standard estimators and test statistics for the analysis of return volatility and covolatility effects on the expected returns and future volatilities. When the volatility effects vanish, one can encounter difficulties that are due to non-identifiability of parameters, or to non-uniform convergence of the objective function used in estimation. Similar problems arise when the second-order causality is examined. Indeed, the null hypotheses of unidirectional second-order causality involve inequality restrictions, which entail identifiability problems of the type considered in this paper (see Gourieroux, Jasiak (2006), Gourieroux (2007), for the definition of causality hypotheses in terms of parameter restrictions).

Appendix 1

Positivity condition

Let us consider a linear form in symmetric positive semi-definite (2,2) matrices:

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \longrightarrow h(\Sigma) = a\sigma_{11} + 2b\sigma_{12} + c\sigma_{22}.$$

This linear form can be equivalently written as:

$$h(\Sigma) = Tr[A\Sigma],$$

where $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and Tr is the trace operator, which computes the sum of diagonal elements of a square matrix.

Lemma 1: The linear form takes nonnegative values for any positive semi-definite matrix Σ, if and only if, matrix A is positive semi-definite.

Proof

Since the set of symmetric positive semi-definite matrices is a positive convex cone, it is equivalent to check the positivity condition on the boundary of the set. This boundary corresponds to the non invertible Σ matrices. These matrices can be written as

$$\Sigma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} (\alpha \ \beta) = \begin{pmatrix} \alpha^2 & \alpha\beta \\ \alpha\beta & \beta^2 \end{pmatrix}.$$

We get

$$h(\Sigma) = a\alpha^2 + 2b\alpha\beta + c\beta^2 \geq 0, \ \ \forall \alpha, \beta.$$

Let us assume $\alpha \neq 0$. The condition becomes:

$$a + 2b(\beta/\alpha) + c(\beta/\alpha)^2 \geq 0, \ \ \forall \alpha, \beta,$$

which is equivalent to $b^2 - ac \leq 0$ (the discriminant of the polynomial of degree 2 is nonpositive), and $a \geq 0$.

By considering the other case $\alpha = 0$, we see that $c \geq 0$.

The set of conditions: $a \geq 0, c \geq 0, ac - b^2 \geq 0$ is exactly the set of conditions for positive semi-definiteness of matrix A. QED

For dimension n larger than 2, it is known that the linear form $\Sigma \to Tr(A\Sigma)$ takes nonnegative values for any positive semi-definite matrix Σ, when A is symmetric positive semi-definite. However, this condition is no longer necessary.

From the proof of Lemma 1, the positivity semi-definiteness condition on A is also required if the linear form has to be nonnegative for any degenerate positive matrix Σ. This is important in ARCH modeling where the realized volatility matrix is generally approximated by squared returns $\Sigma_t = \begin{pmatrix} r_{1t}^2 & r_{1t}r_{2t} \\ r_{1t}r_{2t} & r_{2t}^2 \end{pmatrix}$, that has rank 1. Thus, it is not necessary for us to assume that Σ_t is invertible and to average square returns over a fixed window, for this reason (as suggested, for instance, in Tse, Tsui (2002)).

Appendix 2

Proof of Proposition 1

a) Some Invariance Properties

Proposition A.1 below illustrates invariance properties of $det(A_\infty)$ and $\xi(A_\infty)$ with respect to linear transformations of matrix A_∞.

Proposition A.1 Invariance Properties: For any (n × n) invertible matrices P, Q, we have:

i) $det(PA_\infty Q) = det(P)det(A_\infty)det(Q)$;

ii) $\xi(PA_\infty P') = \xi(A_\infty)$.

Proof: The proof is based on a succession of Lemmas

Lemma 2: If P and Q are (n,n) invertible matrices, we get:

$$cof(PAQ) = det(P)det(Q)Q^{-1}cof(A)P^{-1}.$$

Proof

From the identity $A\,cof(A) = det(A)\,Id$, it follows that

$$(PAQ)Q^{-1}cof(A)P^{-1}det(P)det(Q) = det(A)det(P)det(Q)\,Id = det(PAQ)\,Id.$$

The result follows.

QED

Lemma 3: There exists a (n,n) permutation matrix Δ such that $vec(A') = \Delta vec(A)$. This matrix satisfies $\Delta = \Delta' = \Delta^2$.

Lemma 4: i) $vec(PA) = diag(P)vec(A)$, where diag(P) denotes the bloc-diagonal matrix, with diagonal block P.

ii) $vec(AQ) = \Delta diag(Q')\Delta vec(A)$.

iii) $vec(PAQ) = diag(P)\Delta diag(Q')\Delta vec(A)$.

Proof

i) We have

$$\begin{aligned} PA &= P(a_1, ..., a_n) \text{ (where } a_j \text{ denotes the } j^{th} \text{ column of} A) \\ &= (Pa_1, ..., Pa_n). \end{aligned}$$

Thus, $vec(PA) = \begin{pmatrix} Pa_1 \\ \vdots \\ Pa_n \end{pmatrix} = diag(P)vecA$.

ii) We have

$$\begin{aligned} vec(AQ) &= \Delta vec[(AQ)'] \text{ (by Lemma 3)} \\ &= \Delta vec(Q'A') \\ &= \Delta diag(Q')vec(A') \text{ (from part } i)) \\ &= \Delta diag(Q')\Delta vecA \text{ (by Lemma 3).} \end{aligned}$$

iii) This follows directly from parts i) and ii).

QED

Lemma 5: For any (n,n) invertible matrix P, we have $\xi(PA_\infty P) = \xi(A_\infty)$.

Proof:

Let us consider the transformation:

$$A_\infty \longrightarrow A_\infty^* = PA_\infty Q,$$

where P and Q are deterministic (n,n) invertible matrices. We have:

$$\begin{aligned} vec(A_\infty^*) &= diag(P)\Delta diag(Q')\Delta vec(A_\infty) \text{ (by Lemma 4),} \\ \Omega^* &= Var[vec(A_\infty^*)] = diag(P)\Delta diag(Q')\Delta\Omega\Delta diag(Q)\Delta diag(P'), \\ det(A_\infty^*) &= det(P)det(Q)det(A_\infty), \\ cof(A_\infty^*) &= det(P)det(Q)Q^{-1}cof(A_\infty)P^{-1}, \\ vec[cof(A_\infty^*)] &= det(P)det(Q)diag(Q^{-1})\Delta diag[(P')^{-1}]\Delta vec[cof(A_\infty)]. \end{aligned}$$

If $detP\, detQ > 0$, we find that:

$$\xi(A^*_\infty) = \frac{det(A^*_\infty)}{\sqrt{vec[cof(A^*_\infty)]'\Omega^* vec[cof(A^*_\infty)]}} = \frac{det(A_\infty)}{B_\infty},$$

where

$$\begin{aligned} B_\infty &= vec[cof(A_\infty)]'\Delta diag[(P)^{-1}]\Delta diag[(Q')^{-1}] diagP \Delta diag(Q')\Delta\Omega\Delta diag(Q)\Delta diag(P') \\ & \quad diag(Q^{-1})\Delta diag[(P')^{-1}]\Delta vec[cof(A_\infty)]. \end{aligned}$$

It follows directly that, if $Q = P'$, we have

$$detP\, detQ = (detP)^2 > 0,$$

$B_\infty = vec[cof(A_\infty)]'\Omega vec[cof(A_\infty)]$ and $\xi(PA_\infty P) = \xi(A_\infty)$. The result follows.

b) **Proof of Proposition 1**

We use the invariance properties of $\xi(A_\infty)$ and $det(A_\infty)$ to show Proposition 1.

i) Let us consider a matrix $P = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$. We get:

$$PA_\infty P' = \begin{pmatrix} a_\infty \lambda^2 & b_\infty \lambda\mu \\ b_\infty \lambda\mu & c_\infty \mu^2 \end{pmatrix}.$$

Thus, it is always possible to standardize a_∞ and c_∞ to get $V(a_\infty) = V(c_\infty) = 1$.

ii) Let us now prove that we can find a linear transformation in order to have

$$Cov(a_\infty, b_\infty) = Cov(c_\infty, b_\infty) = 0.$$

For $P = \begin{pmatrix} 1 & \alpha \\ \beta & 1 \end{pmatrix}$, the matrix $A^* = PAP'$ is such that

$$\begin{aligned} a^*_\infty &= a_\infty + 2b_\infty\alpha + c_\infty\alpha^2, \\ b^*_\infty &= a_\infty\beta + (1+\alpha\beta)b_\infty + c_\infty\alpha, \\ c^*_\infty &= a_\infty\beta^2 + 2b_\infty\beta + c_\infty. \end{aligned}$$

The condition $Cov(b^*_\infty, c^*_\infty) = 0$ implies

$$\alpha = -\frac{Cov(a_\infty\beta + b_\infty, a_\infty\beta^2 + 2b_\infty\beta + c_\infty)}{Cov(b_\infty\beta + c_\infty, a_\infty\beta^2 + 2b_\infty\beta + c_\infty)}.$$

By substituting this expression for α in the condition $Cov(a^*_\infty, b^*_\infty) = 0$, we get a polynomial in β of degree 5 (almost surely). This polynomial has at least one real root, which needs to be selected in order to obtain zero covariances.

Appendix 3

The solution in β is finite

When β tends to infinity, the quantity

$$\mu_T(\beta) = < vech\hat{A}_T, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >^2 / < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >$$

tends to $\hat{c}_T^2$. Moreover, the condition $\mu_T(\beta) > \hat{c}_T^2$ is equivalent to:

$$< vech\hat{A}_T, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >^2 - \hat{c}_T^2 < \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix}, \begin{pmatrix} 1 \\ \beta \\ \beta^2 \end{pmatrix} >> 0.$$

It is satisfied for a finite beta value, since the left-hand side of the inequality is a polynomial of degree 3.

Table 1: Critical Values of the Wald Test Statistic for Positive and Negative ϵ

ρ	γ	ϵ Positive			ϵ Negative		
		10	5	1	10	5	1
0.0	0.5	0.945	1.092	1.428	0.903	1.074	1.399
0.0	1.0	0.989	1.141	1.472	0.922	1.092	1.491
0.0	1.5	0.935	1.074	1.359	0.898	1.048	1.423
0.0	2.0	0.896	1.044	1.336	0.871	1.014	1.337
0.1	0.5	0.942	1.081	1.412	0.930	1.064	1.393
0.1	1.0	0.980	1.135	1.465	0.927	1.096	1.457
0.1	1.5	0.936	1.083	1.356	0.894	1.046	1.424
0.1	2.0	0.894	1.046	1.318	0.869	1.015	1.326
0.2	0.5	0.929	1.098	1.411	0.941	1.066	1.384
0.2	1.0	0.975	1.125	1.420	0.946	1.089	1.480
0.2	1.5	0.930	1.076	1.335	0.905	1.037	1.416
0.2	2.0	0.890	1.046	1.309	0.868	1.016	1.331
0.3	0.5	0.928	1.054	1.391	0.930	1.074	1.361
0.3	1.0	0.955	1.119	1.404	0.959	1.102	1.470
0.3	1.5	0.923	1.064	1.306	0.919	1.044	1.396
0.3	2.0	0.896	1.038	1.294	0.876	1.002	1.336
0.4	0.5	0.915	1.062	1.407	0.926	1.079	1.323
0.4	1.0	0.939	1.104	1.380	0.970	1.114	1.441
0.4	1.5	0.904	1.063	1.270	0.921	1.052	1.391
0.4	2.0	0.889	1.026	1.287	0.878	1.015	1.340
0.5	0.5	0.910	1.030	1.367	0.927	1.075	1.354
0.5	1.0	0.930	1.075	1.339	0.991	1.125	1.437
0.5	1.5	0.898	1.049	1.272	0.937	1.058	1.380
0.5	2.0	0.866	1.027	1.268	0.880	1.017	1.340
0.6	0.5	0.899	1.047	1.337	0.934	1.095	1.360
0.6	1.0	0.923	1.058	1.300	1.008	1.147	1.440
0.6	1.5	0.889	1.042	1.259	0.951	1.082	1.358
0.6	2.0	0.858	1.018	1.260	0.888	1.038	1.329
0.7	0.5	0.878	1.052	1.281	0.936	1.108	1.353
0.7	1.0	0.885	1.020	1.276	1.019	1.168	1.414
0.7	1.5	0.857	1.028	1.262	0.965	1.106	1.377
0.7	2.0	0.850	1.005	1.263	0.906	1.057	1.318
0.8	0.5	0.866	1.041	1.255	0.939	1.101	1.372
0.8	1.0	0.859	1.008	1.228	1.034	1.196	1.424
0.8	1.5	0.840	1.020	1.268	0.986	1.121	1.387
0.8	0.2	0.833	0.996	1.265	0.923	1.072	1.333
0.9	0.5	0.841	0.970	1.238	0.941	1.094	1.389
0.9	1.0	0.8433	0.992	1.249	1.049	1.189	1.515
0.9	1.5	0.837	0.995	1.268	0.999	1.155	1.415
0.9	2.0	0.833	0.995	1.265	0.938	1.072	1.327

Table 2: Upper Quantiles of the Student Statistic for α and β

ρ	γ	$\eta_\alpha(10\%)$	$\eta_\alpha(5\%)$	$\eta_\alpha(1\%)$	$\eta_\beta(10\%)$	$\eta_\beta(5\%)$	$\eta_\beta(1\%)$
0.000	0.500	1.523	1.880	2.535	0.997	1.175	1.587
0.000	1.000	1.621	1.963	2.614	1.258	1.500	2.069
0.000	1.500	1.656	1.979	2.625	1.318	1.610	2.221
0.000	2.000	1.657	1.973	2.611	1.331	1.630	2.276
0.100	0.500	1.535	1.866	2.533	1.005	1.183	1.604
0.100	1.000	1.625	1.965	2.593	1.270	1.527	2.064
0.100	1.500	1.650	1.971	2.607	1.327	1.608	2.235
0.100	2.000	1.658	1.966	2.609	1.333	1.635	2.273
0.200	0.500	1.529	1.857	2.574	1.021	1.202	1.639
0.200	1.000	1.631	1.965	2.634	1.300	1.561	2.080
0.200	1.500	1.645	1.982	2.635	1.374	1.666	2.247
0.200	2.000	1.651	1.974	2.621	1.380	1.691	2.315
0.300	0.500	1.538	1.855	2.562	1.054	1.237	1.668
0.300	1.000	1.640	1.959	2.614	1.335	1.622	2.171
0.300	1.500	1.669	1.998	2.631	1.432	1.748	2.360
0.300	2.000	1.664	1.976	2.612	1.455	1.772	2.398
0.400	0.500	1.540	1.870	2.567	1.094	1.284	1.740
0.400	1.000	1.667	1.987	2.683	1.416	1.737	2.319
0.400	1.500	1.691	2.017	2.673	1.539	1.877	2.477
0.400	2.000	1.691	2.010	2.668	1.566	1.939	2.599
0.500	0.500	1.555	1.866	2.548	1.159	1.359	1.843
0.500	1.000	1.670	1.994	2.648	1.537	1.869	2.522
0.500	1.500	1.710	2.025	2.697	1.681	2.047	2.732
0.500	2.000	1.710	2.020	2.673	1.722	2.135	2.882
0.600	0.500	1.551	1.851	2.505	1.248	1.469	1.998
0.600	1.000	1.697	2.006	2.669	1.727	2.085	2.869
0.600	1.500	1.743	2.043	2.721	1.920	2.300	3.131
0.600	2.000	1.724	2.017	2.687	1.982	2.426	3.290
0.700	0.500	1.565	1.901	2.503	1.392	1.641	2.254
0.700	1.000	1.714	2.032	2.691	1.974	2.374	3.297
0.700	1.500	1.741	2.065	2.728	2.252	2.701	3.682
0.700	2.000	1.723	2.054	2.688	2.350	2.855	3.793
0.800	0.500	1.573	1.885	2.516	1.639	1.947	2.670
0.800	1.000	1.726	2.037	2.660	2.444	2.935	4.053
0.800	1.500	1.737	2.074	2.694	2.827	3.359	4.622
0.800	2.000	1.730	2.054	2.700	2.956	3.600	4.834
0.900	0.500	1.607	1.904	2.556	2.281	2.697	3.666
0.900	1.000	1.729	2.049	2.668	3.524	4.200	5.655
0.900	1.500	1.731	2.028	2.694	4.044	4.843	6.567
0.900	2.000	1.739	2.008	2.645	4.291	5.152	6.917

Table 3: Lower Quantiles of the Student Statistic for α

ρ	γ	$\eta_\alpha(1\%)$	$\eta_\alpha(5\%)$	$\eta_\alpha(10\%)$
0.000	0.500	-2.250	-1.510	-1.103
0.000	1.000	-2.283	-1.606	-1.248
0.000	1.500	-2.298	-1.615	-1.263
0.000	2.000	-2.312	-1.636	-1.264
0.100	0.500	-2.263	-1.518	-1.109
0.100	1.000	-2.313	-1.608	-1.239
0.100	1.500	-2.312	-1.642	-1.264
0.100	2.000	-2.327	-1.639	-1.266
0.200	0.500	-2.248	-1.520	-1.122
0.200	1.000	-2.314	-1.622	-1.234
0.200	1.500	-2.334	-1.635	-1.265
0.200	2.000	-2.327	-1.627	-1.262
0.300	0.500	-2.230	-1.528	-1.146
0.300	1.000	-2.267	-1.630	-1.260
0.300	1.500	-2.296	-1.652	-1.279
0.300	2.000	-2.326	-1.654	-1.281
0.400	0.500	-2.225	-1.534	-1.159
0.400	1.000	-2.290	-1.648	-1.279
0.400	1.500	-2.331	-1.686	-1.290
0.400	2.000	-2.317	-1.679	-1.300
0.500	0.500	-2.231	-1.540	-1.157
0.500	1.000	-2.309	-1.663	-1.281
0.500	1.500	-2.394	-1.691	-1.301
0.500	2.000	-2.344	-1.698	-1.308
0.600	0.500	-2.203	-1.525	-1.170
0.600	1.000	-2.352	-1.681	-1.285
0.600	1.500	-2.416	-1.727	-1.317
0.600	2.000	-2.391	-1.722	-1.316
0.700	0.500	-2.218	-1.567	-1.215
0.700	1.000	-2.461	-1.711	-1.316
0.700	1.500	-2.506	-1.757	-1.335
0.700	2.000	-2.468	-1.744	-1.320
0.800	0.500	-2.241	-1.576	-1.238
0.800	1.000	-2.435	-1.737	-1.331
0.800	1.500	-2.510	-1.755	-1.360
0.800	2.000	-2.413	-1.747	-1.342
0.900	0.500	-2.265	-1.614	-1.242
0.900	1.000	-2.423	-1.736	-1.365
0.900	1.500	-2.427	-1.734	-1.331
0.900	2.000	-2.377	-1.752	-1.317

Figure 1: Distribution of beta_infinity

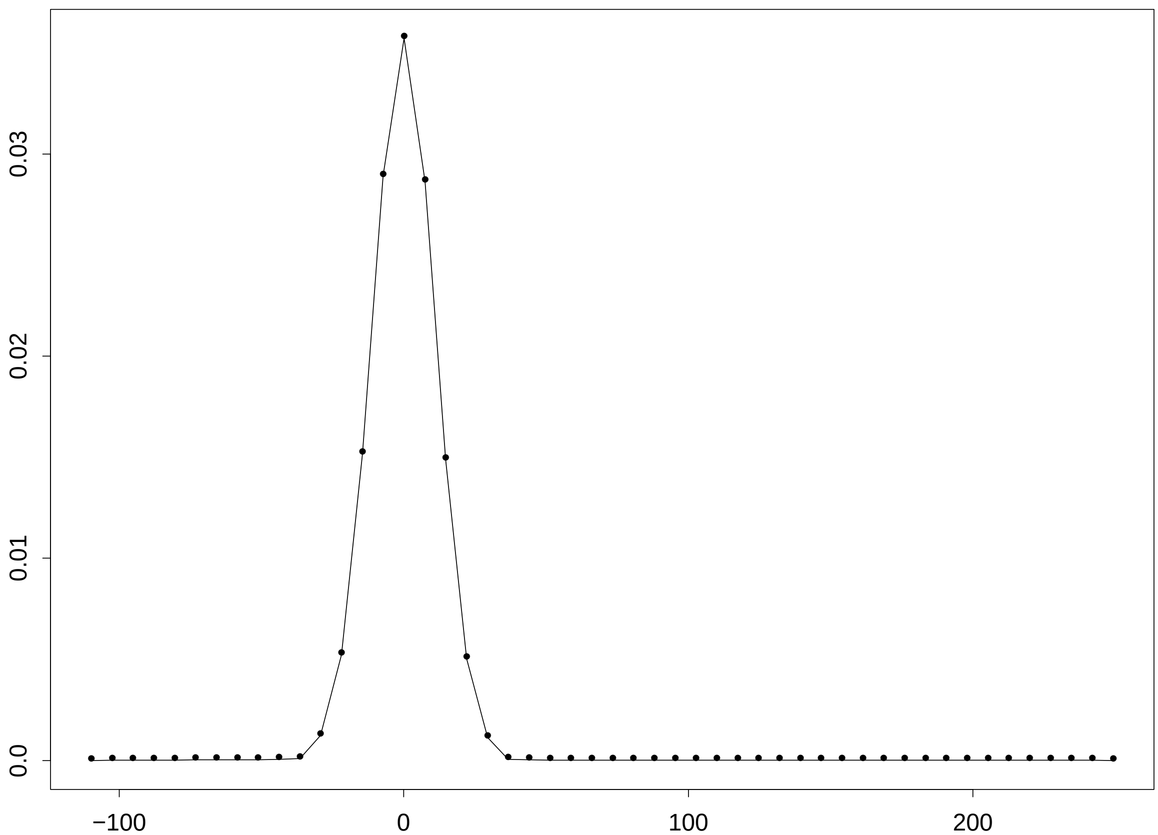


Figure 2a: Finite Sample Distribution of xi_T, Regression 1

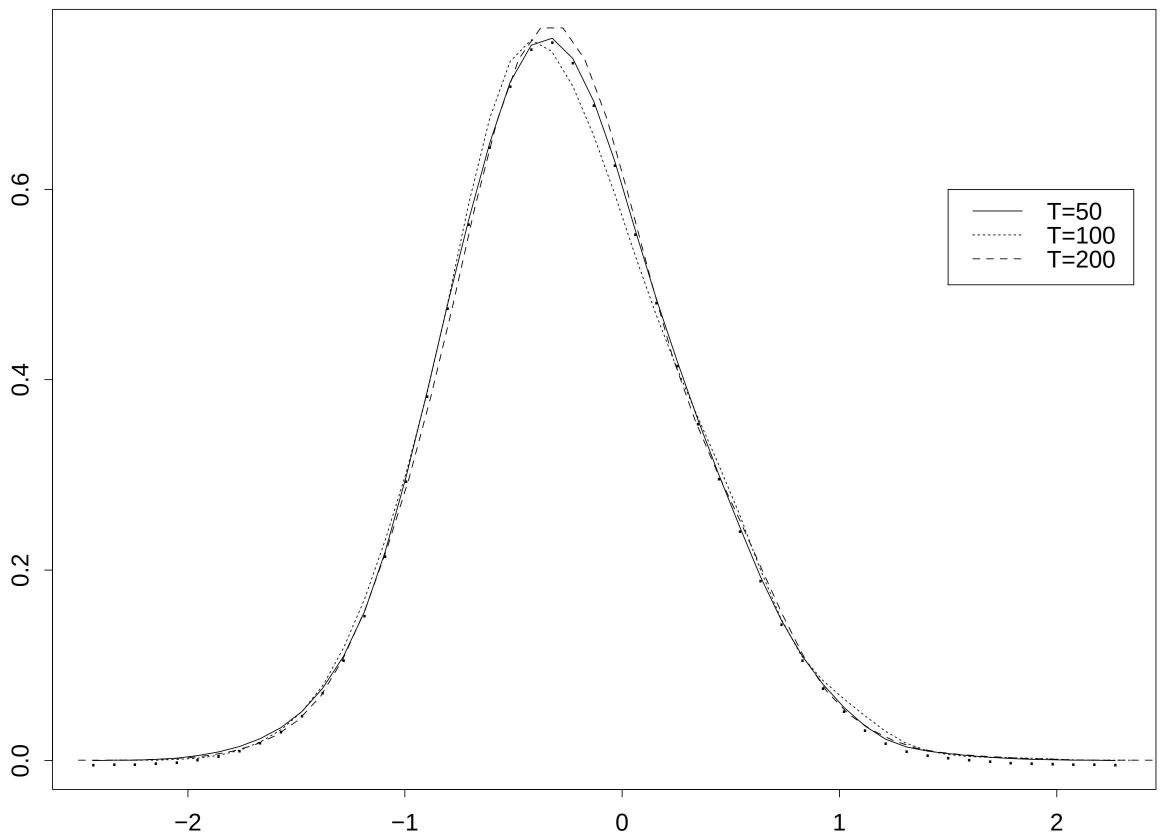


Figure 2b: Finite Sample Distribution of xi_T, Regression 2

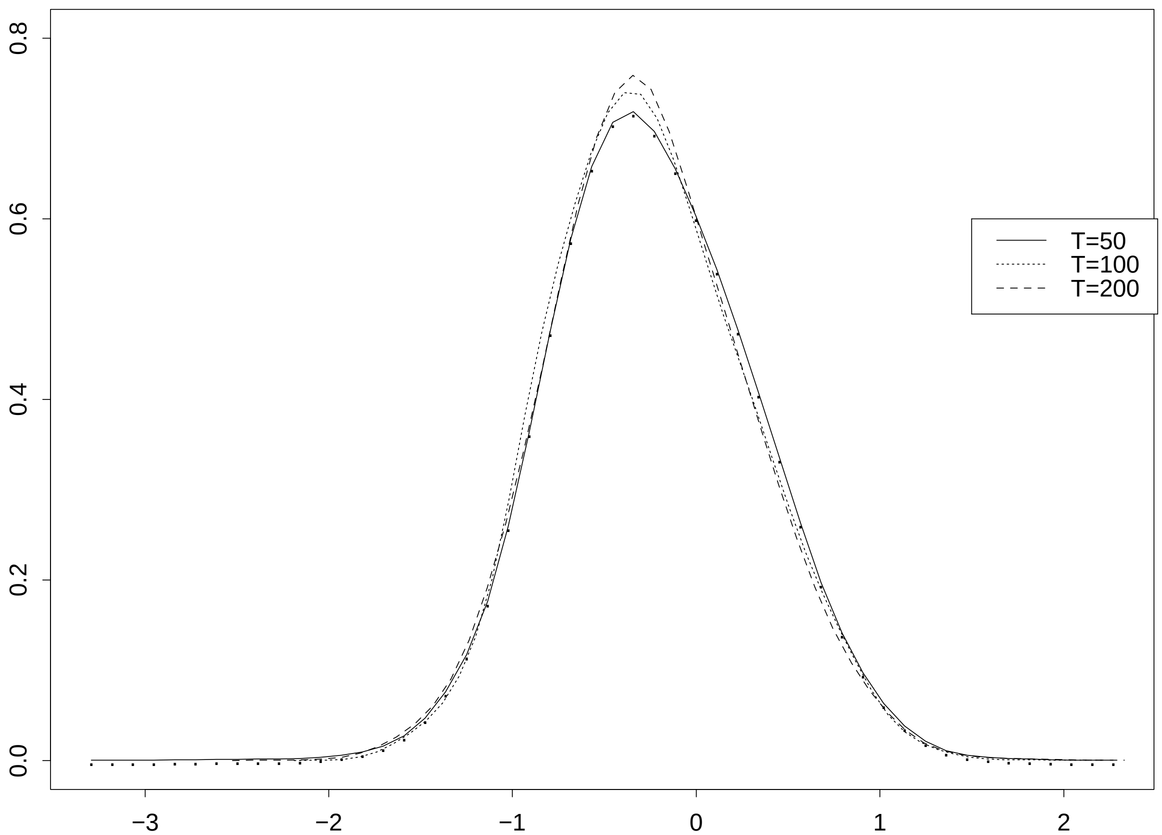


Figure 3a: Finite Sample Distribution of eta(alpha)_T, Reg.1

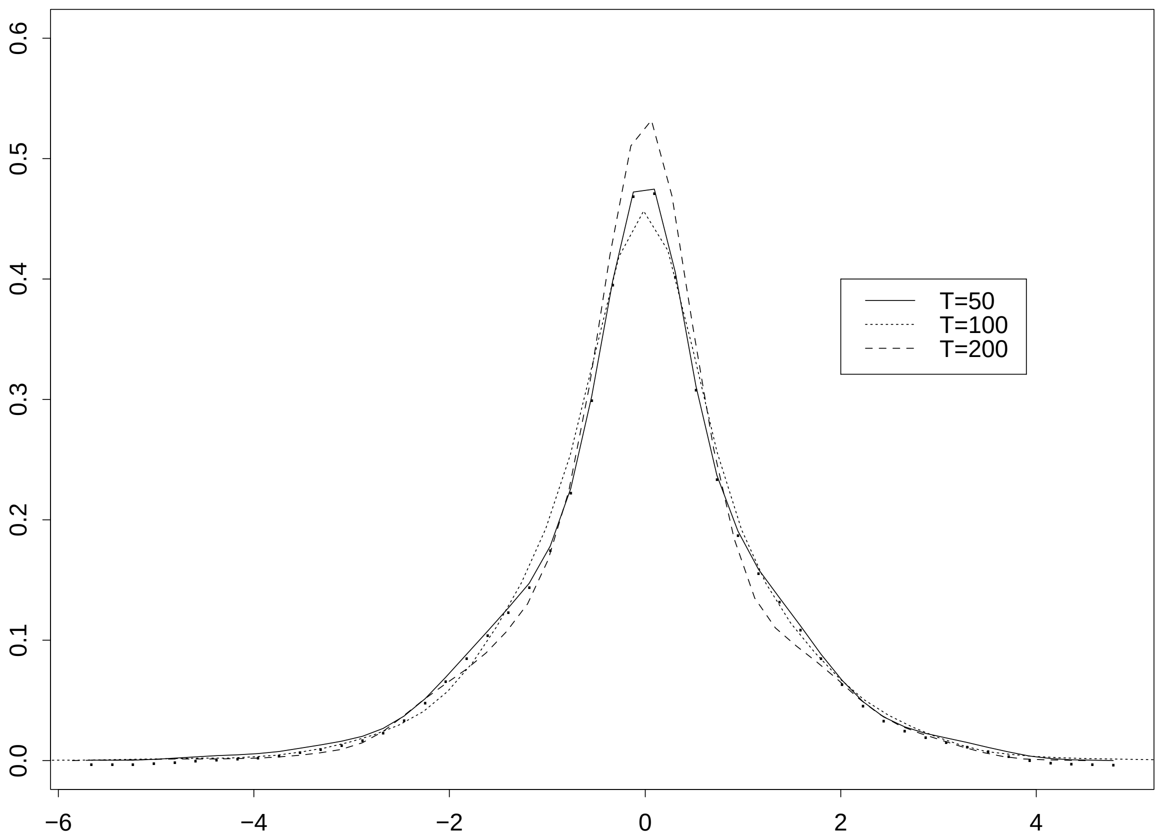


Figure 3b: Finite Sample Distribution of eta(alpha)_T, Reg.2

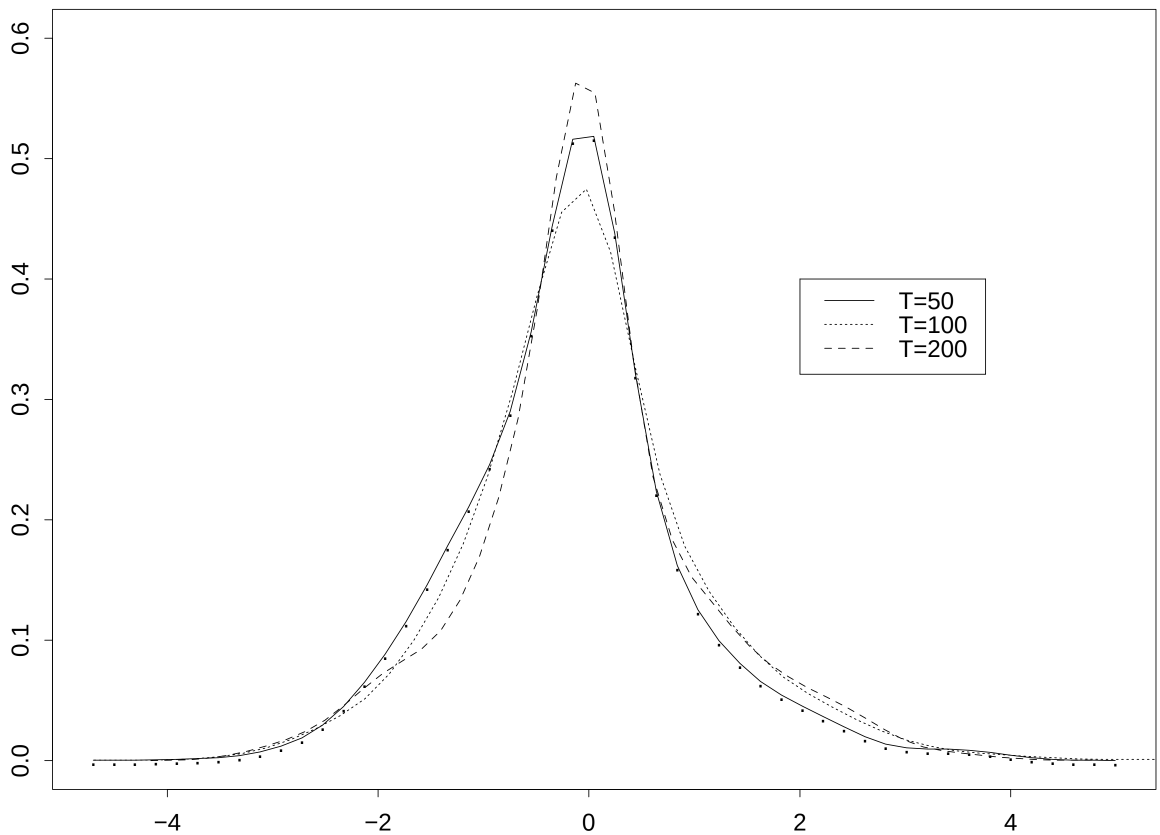


Figure 4a: Finite Sample Distribution of eta(beta)_T, Reg.1

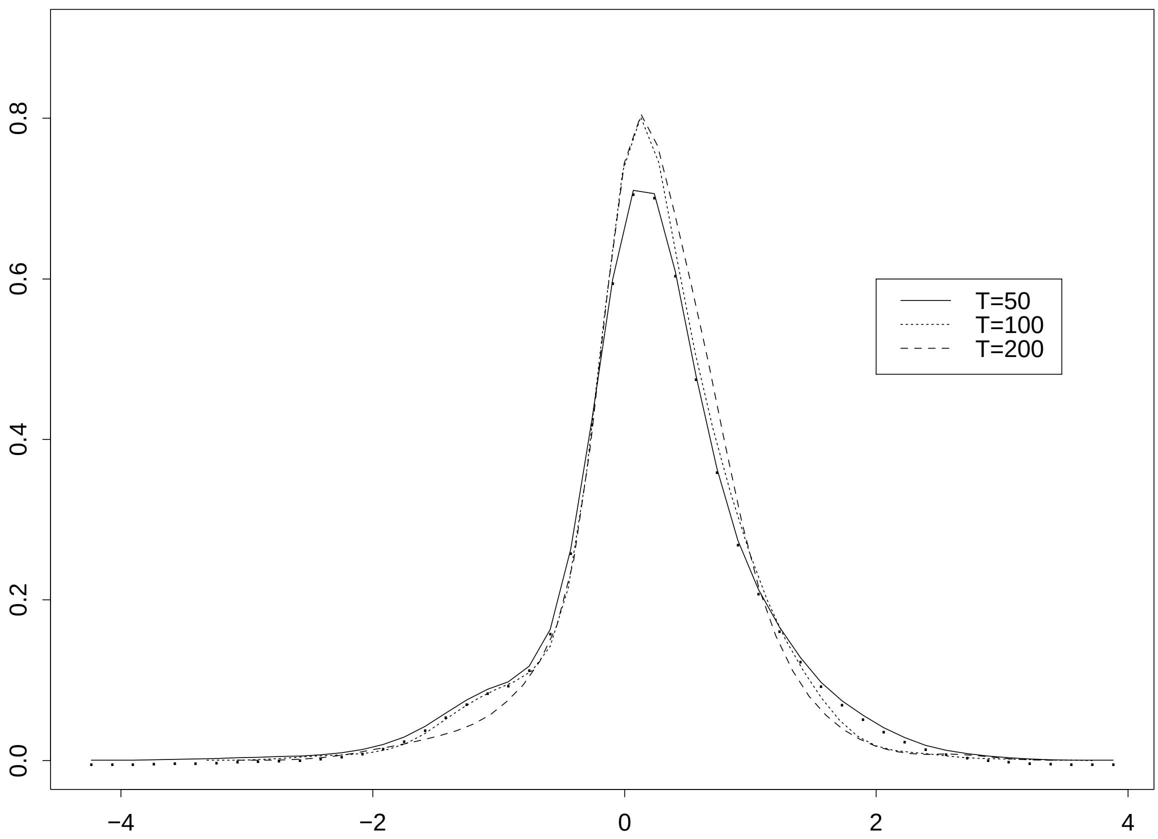
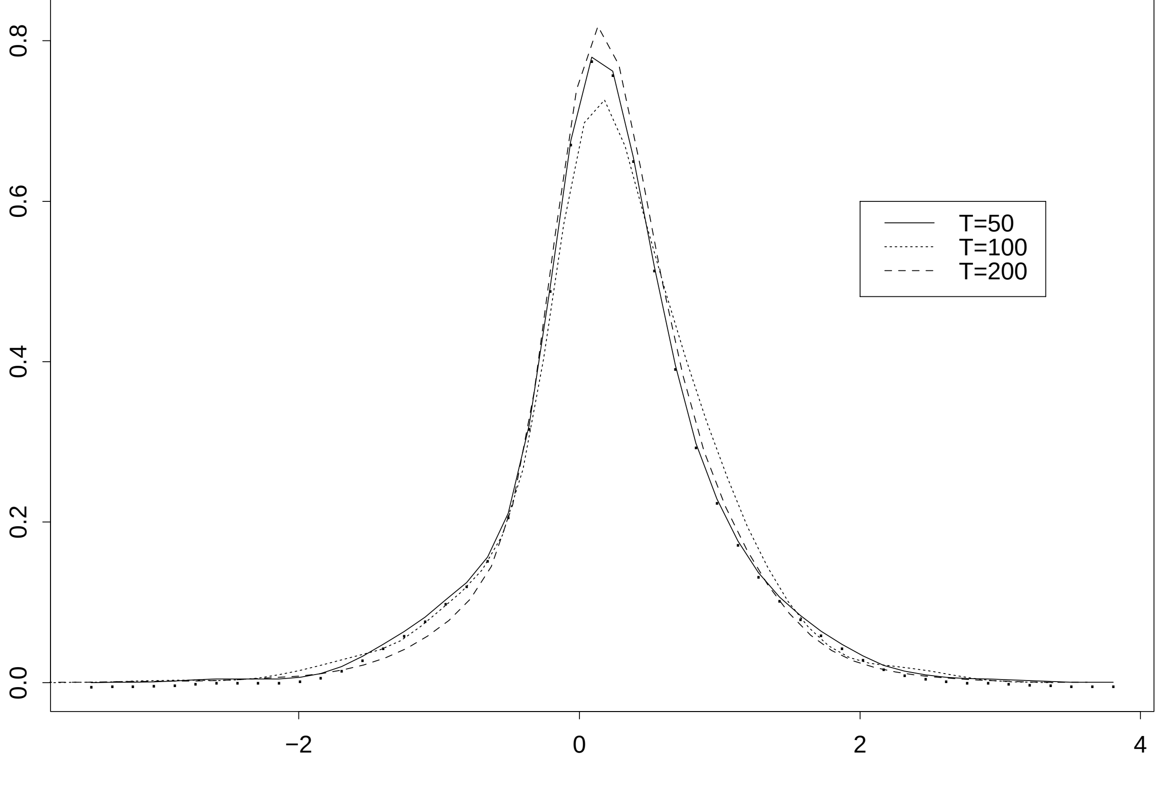


Figure 4b: Finite Sample Distribution of eta(beta)_T, Reg.2



References

[1] Anderson, T. (1984) *An Introduction to Multivariate Statistical Analysis*, 2nd ed. Wiley, New York.

[2] Andrews, D. (1993) Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica* **61**, 821-856.

[3] Andrews, D. (1996) Admissibility of the Likelihood Ratio Test when the Parameter Space is Restricted Under the Alternative. *Econometrica* **64**, 705-718.

[4] Andrews, D. (1999) Estimation when a Parameter is on a Boundary. *Econometrica* **67**, 1341-1383.

[5] Andrews, D. (2001) Testing when a Parameter is on the Boundary of the Maintained Hypothesis. *Econometrica* **69**, 683-734.

[6] Arnott, R. & P. Bernstein (2002) What Risk Premium is Normal? *Financial Analysts Journal* **58**, 64-85.

[7] Arnott, R. & R. Ryan (2001) The Death of the Risk Premium. *Journal of Portfolio Management* **27**, 61-74.

[8] Baba, Y., Engle, R., Kraft, D. & K. Kroner (1990) Multivariate Simultaneous Generalized ARCH Models, DP, University of California, San Diego.

[9] Bilodeau, M. & D. Brenner (1999) *Theory of Multivariate Statistics*. Springer-Verlag.

[10] Bollerslev, T., Chou, R. & K. Kroner (1992) ARCH Modelling in Finance. *Journal of Econometrics* **52**, 5-59.

[11] Bollerslev, T., Engle, R. & J. Wooldridge (1988) A Capital Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy* **96**, 116-131.

[12] Boudoukh, J., Richardson, M. & T. Smith (1993) Is the Ex-Ante Risk Premium Always Positive? A New Approach to Testing Conditional Asset Pricing Models. *Journal of Financial Economics* **34**, 541-553.

[13] Chen, L., Guo, H. & L. Zhang (2006) *Equity Market Volatility and Expected Risk Premium*. Federal Reserve Bank of St. Louis.

[14] Comte, F. & O. Lieberman (2003) Asymptotic Theory for Multivariate GARCH Processes. *Journal of Multivariate Analysis* **84**, 61-84.

[15] Cragg, J., and S. Donald (1996): "On the Asymptotic Properties of LDU-Based Tests of the Rank of a Matrix", JASA, 91, 1301-1309.

[16] Cragg, J., and S. Donald (1997): "Inferring the Rank of a Matrix", Journal of Econometrics, 76, 223-250.

[17] Domowitz, I. & C. Hakkio (1985) Conditional Variance and the Risk Premium in the Foreign Exchange Market. *Journal of International Economics* **22**, 301-315.

[18] Engle, R., Granger, C. & D. Kraft (1984) Combining Competing Forecasts of Inflation Using A Bivariate ARCH Model. *Journal of Economic Dynamics and Control* **8**, 151-165.

[19] Engle, R. & K. Kroner (1995) Multivariate Simultaneous Generalized ARCH. *Econometric Theory* **11**, 122-150.

[20] Francq, C. & J.H. Zakoian (2006) Quasi-Likelihood Inference in GARCH Processes when Some Coefficients are Equal to Zero. CREST, DP.

[21] Gill, L., and A. Lewbel (1992): "Testing the Rank and Definiteness of Estimated Matrices with Applications to Factor, State-Space and ARMA Models", JASA, 87, 766-776.

[22] Gourieroux, C. (2007) Positivity Condition for a Bivariate Autoregressive Volatility Specification. *Journal of Financial Econometrics* **5**, 624-636.

[23] Gourieroux, C., Holly, A. & A. Monfort (1980) Kuhn-Tucker, Likelihood Ratio and Wald Tests for Nonlinear Models with Inequality Constraints on the Parameters. Harvard D.P. 770.

[24] Gourieroux, C., Holly, A. & A. Monfort (1982) Likelihood Ratio Test, Wald Test and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters. *Econometrica* **50**, 63-80.

[25] Gourieroux, C. & J. Jasiak (2006) Nonlinear Causality, with Application to Liquidity and Stochastic Volatility. CREST Working Paper.

[26] Gourieroux, C., Jasiak, J. & R. Sufana (in press) The Wishart Autoregressive Process of Multivariate Stochastic Volatility *Journal of Econometrics.*

[27] Gourieroux, C. & A. Monfort (1989) *Statistique et Modeles Econometriques*, Vol 2, Paris, Economica. English translation (1995) *Statistics and Econometric Models*, Vol 2, Cambridge Univ. Press, Cambridge UK.

[28] Gourieroux, C., Monfort, A. & E. Renault (1995) Inference in Factor Models. In Maddala G., Phillips, P.C.B., and T. Srinivasan (eds) *Advances in Econometrics and Quantitative Economics.* Oxford: Brockwell, 311-353.

[29] Guggenberger, P., Hahn, J. & K. Kim (2006) Specification Testing Under Moment Inequalities. UCLA DP.

[30] Hakkio, C. & A. Sibert (1995) The Foreign Exchange Risk Premium: Is it Real? *Journal of Money, Credit and Banking* **27**, 301-317.

[31] Hansen, P. (2003) Asymptotic Tests of Composite Hypotheses. Brown University, DP 2003/09.

[32] Iglesias, E. & G. Phillips (2005) Bivariate ARCH Models: Finite Sample Properties of QML Estimators and an Application to an LM-type Tests. *Econometric Theory* **21**, 1058-1086.

[33] Jeantheau, T. (1998) Strong Consistency of Estimators for Multivariate ARCH Models. *Econometric Theory* **14**, 70-86.

[34] Kodde, D. & F. Palm (1986) Wald Criteria for Jointly Testing Equality and Inequality Restrictions. *Econometrica* **54**, 1243-1248.

[35] Macklem, T. (1991) Forward Exchange Rates and Risk Premium in Artificial Economies. *Journal of International Money and Finance* **10**, 365-391.

[36] Moon, R. & F. Schorfheide (2005) Empirical Likelihood Estimation with Inequality Moment Restrictions. University of Pennsylvania D.P.

[37] Ostdiek, B. (1998) The World Ex-Ante Risk Premium: An Empirical Investigation. *Journal of International Money and Finance* **17**, 967-999.

[38] Robin, J. and R. Smith (2000): "Tests of Rank", Econometric Theory, 16, 151-175.

[39] Silberberg, G. & S. Pafka (2001) A Sufficient Condition for the Positive Definiteness of the Covariance Matrix of the Multivariate GARCH Model. CEU, 7/2001.

[40] Tse, Y. & A. Tsui (2002) A Multivariate GARCH Model with Time Varying Correlations. *Journal of Business and Economic Statistics* **20**, 351-362.

[41] Walsh, K. (2006) Is the Ex-Ante Risk Premium Always Positive? Further Evidence. *Australian Journal of Management* **31**, 93-113.

[42] Wolak, F (1991) The Local and Global Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models. *Econometrica* **59**, 981-995.