INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2006-30

A Degeneracy in the Analysis of Volatility and Covolatility Effects

C. GOURIEROUX[1]
J. JASIAK[2]

1 CREST, CEPREMAP and University of Toronto, Canada: gouriero@ensae.fr
2 York University, Canada. Email : jasiakj@yorku.ca

A Degeneracy in the Analysis of Volatility and Covolatility Effects

Christian Gourieroux*

Joann Jasiak†

This version: December 19, 2006

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

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

Abstract

A Degeneracy in the Analysis of Volatility and Covolatility Effects

Numerous financial applications examine the risk effect on future expected returns and volatilities. This effect is generally specified as a linear function of lagged volatilities and covolatilities. The aim of this paper is to point out that, in this framework, estimation and testing need to be carried out with caution as some standard regularity conditions of asymptotic theory are not satisfied. This paper shows the adjustements in the asymptotic distributions of selected frequently used test statistics and estimators.

Keywords: Volatility Model, Risk Premium, BEKK Model, Volatility Transmission, Identifiability, Boundary, Invertibility Test.
JEL number: C22, C32, G10, G12.

1 Introduction

One of the main financial problems is the measurement of risks and the analysis of risk effects on the distribution of future returns. In many financial models, the risks are represented by a volatility-covolatility matrix and the effect of risk on expected returns or future volatilities is specified as an affine function of the current and lagged realized volatilities and covolatilities. Let us consider a 2-asset framework with the 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:

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

for example. This model can be used to examine the size of the ex-ante equity risk premium, its sign, and its existence. Boudoukh et al (1993), Ostdiek (1985), Arnott, Ryan (2001), Arnott, Bernstein (2002), Chen, Guo, Zhang (2006), Walsh (2006) have shown that the risk premium can take either a positive or a negative sign depending on the environment, and tested the positivity of the conditional risk premium using instrumental variables. Moreover, under the CAPM framework, there exists a relation between the expected return and the variance of the market portfolio. As a consequence, the effect of the variances-covariances of the assets is captured by a single market portfolio. This implies that the risk premium is of reduced rank.

An analogous problem concerning the sign and rank of the risk premium arises in foreign exchange markets [see e.g. Domowitz, Hakkio (1985), Macklem (1991) , Hakkio, Sibert (1995)]. The sign of the foreign exchange real risk premium can depend on the ratio of volatilities in both countries. Moreover, the test of zero risk premium is equivalent to a test of the null hypothesis that the forward exchange rate is an unbiased predictor of the future spot exchange rate.

Similar specifications of volatility-covolatility effects are introduced in multivariate ARCH models (see, e.g. Engle, Granger, Kraft (1984), Bollerslev, Engle, Wooldridge (1988), Bollerslev, Chou, Kroner (1992)) to describe the expected future volatility. For instance, a so-called vech-representation implies:

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

where $\tilde{\sigma}_{ij,t} = r_{i,t} r_{j,t},\ i,j = 1,2.$

In this model, it is interesting to test for the significance of the effect of lagged realized volatility, or to check if the realized volatility effects can be summarized by a smaller number of factors as in the BEKK model (Baba, Engle, Kraft, Kroner (1990)).

A linear form in the volatilities-covolatilities can always be written 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. In practice, the matrix of sensitivity coefficients A is estimated from T asset returns. We assume that the estimator $\hat{A}_T$ is a consistent, asymptotically normal estimator of A.

This paper considers hypotheses testing and constrained estimation concerning matrix A, mainly for A of dimension 2 × 2. The hypotheses of interest are:

1) the hypothesis of noninvertibility of matrix A;

2) the hypothesis that matrix A is positive semi-definite. Indeed, this condition is equivalent to the nonnegativity of the linear form $Tr(A\Sigma)$ (see Appendix 1, Lemma 1).

The constrained estimation concerns estimation of A assuming that its rank is less or equal to 1.

At a first sight, these tests and estimation problems [1] can seem rather standard. For instance, the invertibility of matrix A is usually based on the singular value decomposition of the (asymptotically) Gaussian random matrix $\hat{A}_T$ [see Anderson (1989), Gourieroux, Monfort, Renault (1995), Bilodeau, Brenner (1999)]. The tests for matrix positivity are based on asymptotic test procedures for the inequality restrictions $ac - b^2 \geq 0$, $a \geq 0$ [see e.g. Gourieroux, Monfort (1989), Wolak (1991)]. The estimation of A under the hypothesis of reduced rank is performed by quasi-maximum likelihood methods as for the BEKK model [Engle, Kroner (1995), Jeantheau (1998), Comte, Lieberman (2000)].

The purpose of this paper is to point out identifiability problems, boundary problems and degeneracies occurring in these frameworks. They imply complicated asymptotic distributions both for estimators and tests, can render standard practices misleading and stylized facts questionable.

[1] Similar problems exist for 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 only sufficient for a larger dimension (Silberberg, Pafka (2001)).

In the paper, we apply standard estimators and test statistics and derive their correct asymptotic distributions. The analysis of the asymptotic admissibility of the test statistics and their potential improvement are out of the scope of the present paper. In Section 2, we consider the Wald test for non invertibility of matrix A based on its estimated determinant det $\hat{A}_T$. We explain why the Wald test statistic has a non-Gaussian distribution in the degenerate case A=0, and study how this distribution depends on the asymptotic variance of the random matrix $\hat{A}_T$. Section 3 discusses the constrained estimation of A when this matrix has reduced rank. The distribution of the constrained estimator is non-standard if A=0. Section 4 considers the test of positive semi-definiteness, that is, of the hypothesis defined by the inequality constraints $a \geq 0, c \geq 0, ac-b^2 \geq 0$. In this case also, the standard asymptotic theory is not valid if A=0. We explain how it can be corrected, when A is unconstrained [respectively, when A is 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.

2 Testing for Invertibility

As mentioned in the introduction , there exist different approaches to testing for noninvertibility of matrix A. The first one consists in performing a singular value decomposition of matrix A and checking if the smallest singular value is close to zero. The distribution of the smallest singular value is rather complicated and depends on the number of zero singular values [see Anderson (1989), Bilodeau, Brenner (1999)]. The second approach consists in testing directly the significance of the determinant. In the remainder of this section, the second approach is considered for the following two reasons. First, the determinant inequality restrictions characterize positive semi-definiteness. Second, the degeneracy problem is easier to explain in this context.

2.1 The constrained and unconstrained models

We are interested in a n $\times$ n matrix parameter A, which can be consistently estimated by an asymptotically Gaussian estimator $\hat{A}_T$. Let us denote by vec A the vector of length n^2 obtained by stacking the columns of matrix A. We assume that:

$$\sqrt{T}[vec(\hat{A}_T) - vec(A)] \overset{d}{\to} N(0, \Omega), \tag{2.1}$$

where Ω is a $(n^2 \times n^2)$ invertible matrix and $\overset{d}{\to}$ denotes the convergence in distribution. $\hat{A}_T$ summarizes the relevant information about A contained in the data. Thus, (2.1) will be considered as the unconstrained (asymptotic) model.

In this section, we want to test the null hypothesis of noninvertibility of matrix A:

$$H_0 : \text{ (A is not invertible) } = (\det A = 0). \tag{2.2}$$

2.2 Wald Test statistic

A standard approach to testing the null hypothesis H_0 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 of elements of A, we get:

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

Then, the Wald test statistic is defined by:

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

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%.

2.3 The degenerate case

The standard approach described above is valid if $vec(cof(A)) \neq 0$, that is, if $A \neq 0$. Otherwise, the asymptotic properties of the Wald test statistic are significantly altered. Indeed, 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 larger;

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

A similar analysis can be done for the test statistic $\hat{\xi}_T$. By noting that $cof(\sqrt{T}\hat{A}_T) = T^{(n-1)/2}cof(\hat{A}_T)$, we see that $\hat{\xi}_T \xrightarrow{d} \xi(A_\infty)$, where

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

The distributions of $det(A_\infty)$ and $\xi(A_\infty)$ are complicated, but feature invariance properties with respect to linear transformations of matrix A_∞ (see Appendix 2).

Proposition 1: For any (n $\times$ n) invertible matrices P, Q, we have:

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

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

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

This degenerate case cannot be disregarded or circumvented [2]since in practice the hypothesis A=0 can have very appealing interpretations. For instance, this hypothesis is to be considered for determining the autoregressive order of a multivariate ARCH model [3]. Also, in the application to risk premium, the constraint A=0 characterizes the hypothesis of nonpredictability of asset returns.

2.4 Consequences for test results

The multiplicity of limiting distributions of the Wald test statistic under the null hypothesis requires a careful analysis of type I error, since the asymptotic similarity on the boundary condition is violated [see Hansen (2003)]. For instance, let us consider the practice of rejecting the null hypothesis if the Wald statistic $\hat{\xi}_T$ is sufficiently large in absolute value. The condition on the (asymptotic) type I error is:

$$sup_{H_0} \lim_{T\to\infty} P[|\hat{\xi}_T| > c] = \alpha$$
$$\iff \quad sup[sup_{A:detA=0,A\neq 0} \lim_{T\to\infty} P(|\hat{\xi}_T| > c), sup_{A=0} \lim_{T\to\infty} P(|\hat{\xi}_T| > c)] = \alpha$$

[2]For example, by introducing null hypotheses indexed by the number T of observations, such as $H_{0,T} : [detA = 0, ||A|| > h(T)]$, where $h(T)$ is strictly positive and tends to zero at an appropriate rate, when T tends to infinity. Such a methodology is followed in the test of switching regimes, for the parameter representing the unknown switching date [Andrews (1993)].

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

$$\begin{aligned}
&\Longleftrightarrow \quad sup[(P(|X| > c), P(|\xi(A_\infty)| > c)] = \alpha \text{ (where } X \sim N(0,1))\\
&\Longleftrightarrow \quad c(\alpha) = Max[\Phi^{-1}(1 - \frac{\alpha}{2}), Q(\alpha, \Omega)], \qquad (2.7)
\end{aligned}$$

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

$$P[|\xi(A_\infty)| > Q(\alpha, \Omega)] = \alpha, \qquad (2.8)$$

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

Despite that function Q has a complicated expression (see Section 2.5), the value $Q(\alpha, \Omega)$ is easily approximated by Monte-Carlo, that is, by the corresponding sample quantile computed from simulated values $\xi(A^s_\infty), s = 1, ..., S$, where $vec(A^s_\infty), s = 1, ...S$, are independently drawn in the Gaussian distribution $N(0, \hat{\Omega}_T)$.

2.5 Is an adjustement necessary in the symmetric (2,2) case?

As mentioned in the introduction, we are especially interested in (2,2) symmetric matrices $\hat{A}_T = \begin{pmatrix} \hat{a}_T & \hat{b}_T \\ \hat{b}_T & \hat{c}_T \end{pmatrix}$ with corresponding limits $A_\infty = \begin{pmatrix} a_\infty & b_\infty \\ b_\infty & c_\infty \end{pmatrix}$. By Proposition 1, we can restrict the choice of Ω while searching for the possible distribution of $\xi(A_\infty)$ [see Appendix 3].

Proposition 2: Up to a transformation $A_\infty \longrightarrow PA_\infty P'$, we can assume that

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

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

Such a Gaussian random matrix can be easily simulated by writing

$$a_\infty = \sqrt{1-\rho^2}X + \epsilon\rho F, ,\ c_\infty = \sqrt{1-\rho^2}Y + \rho F,\ b_\infty = \gamma Z, \qquad (2.9)$$

where X, Y, Z, F are iid standard normals.

The Wald test statistic is:

$$\begin{aligned}
\xi(A_\infty) &= \frac{a_\infty c_\infty - b^2_\infty}{\sqrt{(c_\infty, -2b_\infty, a_\infty)\tilde{\Omega}(c_\infty, -2b_\infty, a_\infty)'}}\\
&= \frac{a_\infty c_\infty - b^2_\infty}{\sqrt{c^2_\infty + a^2_\infty + 2\epsilon\rho^2 c_\infty a_\infty + 4b^2_\infty\gamma^2}}. \qquad (2.10)
\end{aligned}$$

We provide in Table 1 the upper quantiles at 10%, 5% and 1% of the variable $|\xi(A_\infty)|$ for different values of parameters ρ, γ and $\epsilon = +/-1$. These quantiles are computed by Monte-Carlo with 5000 replications. They can be directly compared to the critical values 1.64, 1.96, 2.57 of the standard normal distribution, which corresponds to the case when $detA = 0$ with $A \neq 0$. We observe that these values are systematically smaller than their Gaussian counterparts. This implies that the standard Wald test does not need to be corrected for the degeneracy at A=0.

3 Constrained Estimation of A

3.1 The Example of BEKK model

To ensure the positivity of the volatility $H_t = V_t(r_{t+1})$, the multivariate GARCH literature (Engle, Kroner (1995)) proposed the following constrained specification [4]:

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

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:

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

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. The usual practice consists in optimizing a quasi-likelihood function with respect to parameters M (and N) [see e.g. Engle, Kroner (1995), Comte, Lieberman (2003)]. Let us consider the bidimensional case, $A = \begin{pmatrix} m_1^2 & m_1 m_2 \\ m_1 m_2 & m_2^2 \end{pmatrix}$. 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:

[4] For ease of exposition, we have introduced only 1 positive component by lag.

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

where $\alpha = m_1^2, \beta = m_2/m_1$ (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 for multivariate BEKK models doesn't hold for estimation 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) introduce the identifiability condition $m_1 > 0$ (Proposition 2.1). This condition eliminates both the global and local identifiability problems.

In the next section, we derive the correct asymptotic distributions of the minimum distance estimators of α and β based on a consistent, asymptotically normal estimator of A. For the application to BEKK model, we assume that the quasi-maximum likelihood estimator is asymptotically normal. This requires some additional assumptions for 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).

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 [5].

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

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

The objective function 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. Assumption 3 in Andrews (1999), p. 1349, is not satisfied and new asymptotic results need to be derived.

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

The objective function can be concentrated with respect to α. 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:

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

Appendix 4 shows that this solution is finite.

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 is a root of a polynomial of degree 5.

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, then the Jacobian has 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

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

It follows that $\hat{\beta}_T$ tends to a limit β_∞, which is a solution of the optimization:

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

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 $(\sqrt{T}\hat{\alpha}_T, \hat{\beta}_T) \overset{d}{\to} (\alpha_\infty, \beta_\infty)$, 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 $[J(\hat{\alpha}_T, \hat{\beta}_T)\tilde{\Omega}_T^{-1} J(\hat{\alpha}_T, \hat{\beta}_T)']^{-1}$. If $A \neq 0$, this statistic tends in distribution to a standard normal. If A=0, statistic $\hat{\eta}_T$ 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

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

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 statistics $|\eta^{\alpha}_{\infty}|$ and $|\eta^{\beta}_{\infty}|$, respectively, for Gaussian matrices already considered in Section 2. These quantiles have been derived 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. Moreover both of them are much more sensitive to parameter γ. They differ significantly from the Gaussian quantiles 1.64, 1.96, 2.57, especially for parameter β.

Figure 1 shows the asymptotic 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 Test for positivity

Let us now focus on the test for positivity of the symmetric matrix A. This test depends on the maintained hypothesis, that is either "A unconstrained", or "A of reduced rank". The two cases are discussed below.

4.1 A unconstrained

Usually the null hypothesis is written as $H_0 : \{a \geq 0, ac - b^2 \geq 0\}$, and the test of inequality restrictions is performed along the lines developped [6] by [Gourieroux, Holly, Monfort (1980),

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

(1982), Kodde, Palm (1986), Gourieroux, Monfort (1989), Wolak (1991)]. However, due to the degeneracy problem, this standard technique cannot be applied. The reason is that it requires the Jacobian of the transformations defining the constraints, that is, $(a,b,c) \rightarrow (a, ac-b^2)$ be of full rank on the boundaries of the null hypothesis. However, for A=0, the Jacobian $\begin{pmatrix} 1 & 0 & 0 \\ c & -2b & a \end{pmatrix}$ is of reduced rank.

In other words, the degeneracy problem 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, we deduce from $ac-b^2 \geq 0$ that c (resp a) is nonnegative. 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 asymptotic 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 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 take three different expressions:

i) When $\hat{A}_T >> 0$, it is equal to $\hat{A}_T$.

ii) The solution of the minimization (4.2) can be a positive semi-definite matrix of rank 1.

iii) The solution of the minimization can be 0.

Under standard regularity conditions, the maximal value of the type I error under the null is attained at A=0, and is computed from a mixture of chi-square distributions, with weights corresponding to the probabilities of the three regimes i), ii), iii) computed under A=0.

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

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

Thus, (4.3) defines an asymptotic optimization problem under A=0. The regimes are determined by the possible values of the 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

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

From the standard asymptotic theory for testing inequality constraints [see e.g. Gourieroux, Holly, Monfort (1980), Gourieroux, Monfort (1989), Wolak (1991)], it follows that 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 [7]:

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

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 computed by simulations.

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

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 seen in Section 3, the asymptotic distribution of this 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.

5 Finite Sample Properties

To study the finite sample properties of the standard statistics, we generate iid Gaussian returns $(r_{1,t}, r_{2,t})'$, that are $IIN(0, Id)$. The number of observations is T=50, 100, 200. Then, 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 bivariate risk premium, while the second regression considers the problem of volatility transmission.

For each regression, we provide 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 $\hat{\xi}_T$ are displayed in Figures 2a-2b for the two regressions. We observe fat tails, and different limiting distributions for the two regressions. Indeed, the limiting OLS covariance matrices are different 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 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 correct limiting distributions of standard estimators and test statistics for the analysis of the effect of volatilities and covolatilities on expected returns and future volatilities. The difficulties are due to a lack of identifiability and to the non-uniform convergence of the objective function, when these effects vanish. Similar problems arise when the second-order causality is examined. Indeed, the null hypotheses of unidirectional second-order causality involve inequality restrictions with identifiability problems of the type considered in this paper (see Gourieroux, Jasiak (2006), Gourieroux (2006), for the definition of causality hypotheses in terms of parameter restrictions).

Appendix 1

Positivity condition

Let us consider a linear form defined on 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 that 2, it is known that the linear form $\Sigma \rightarrow Tr(A\Sigma)$ 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 the square 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 our problem to average square returns on a given window to be sure that Σ_t is invertible (as suggested, for instance, in Tse, Tsui (2002)).

Appendix 2

Proof of Proposition 1

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 $Acof(A) = det(A)Id$, 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

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

iii) This is a direct consequence of parts i) and ii).

QED

Let us now consider the transformation:

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

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

$$\begin{aligned} vec(A^*_\infty) &= diag(P)\Delta diag(Q')\Delta vec(A_\infty) \text{ (from Lemma2)}, \\ \Omega^* &= Var[vec(A^*_\infty)] = diag(P)\Delta diag(Q')\Delta\Omega\Delta diag(Q)\Delta diag(P'), \\ det(A^*_\infty) &= det(P)det(Q)det(A_\infty), \\ cof(A^*_\infty) &= det(P)det(Q)Q^{-1}cof(A_\infty)P^{-1}, \\ vec[cof(A^*_\infty)] &= det(P)det(Q)diag(Q^{-1})\Delta diag[(P')^{-1}]\Delta vec[cof(A_\infty)]. \end{aligned}$$

If $detP\, detQ > 0$, we deduce 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

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

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

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

Appendix 3

Proof of Proposition 2

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 chosen in order to obtain zero covariances.

Appendix 4

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 for 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 for 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

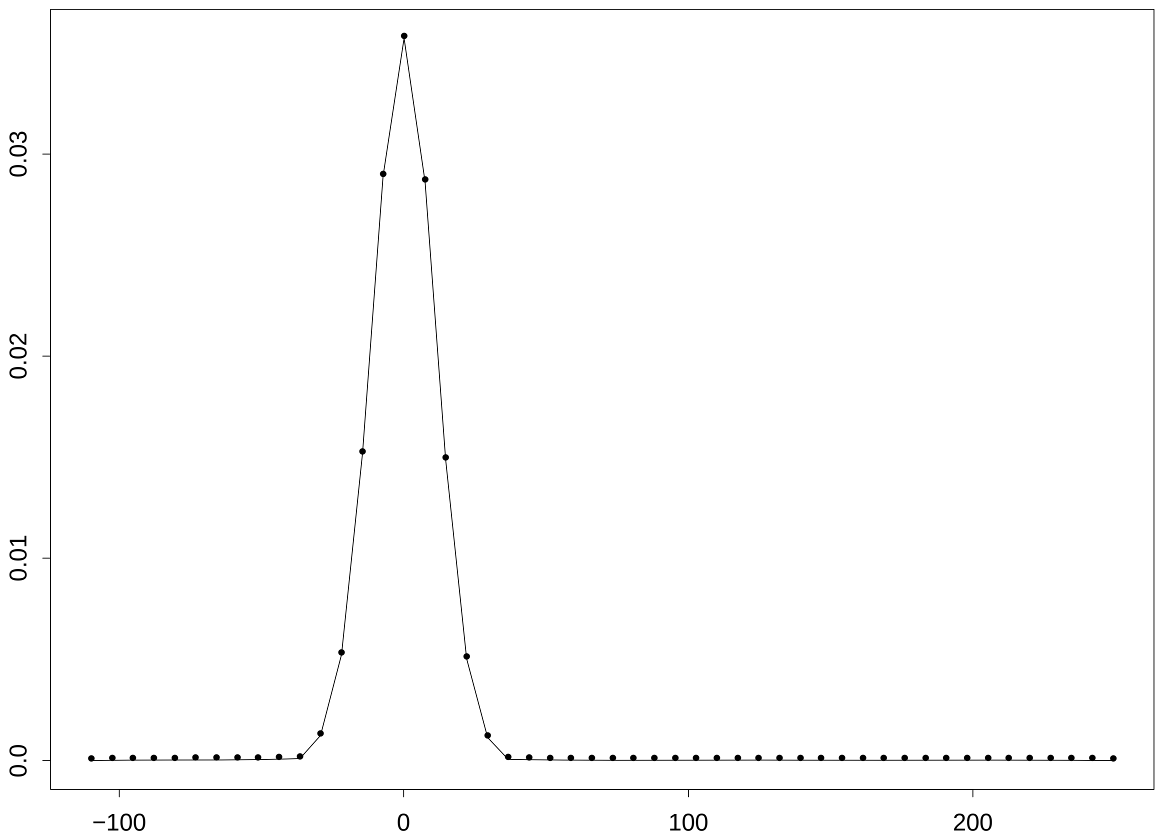


Figure 2a: Finite Sample Distribution of xi_T, Regression 1

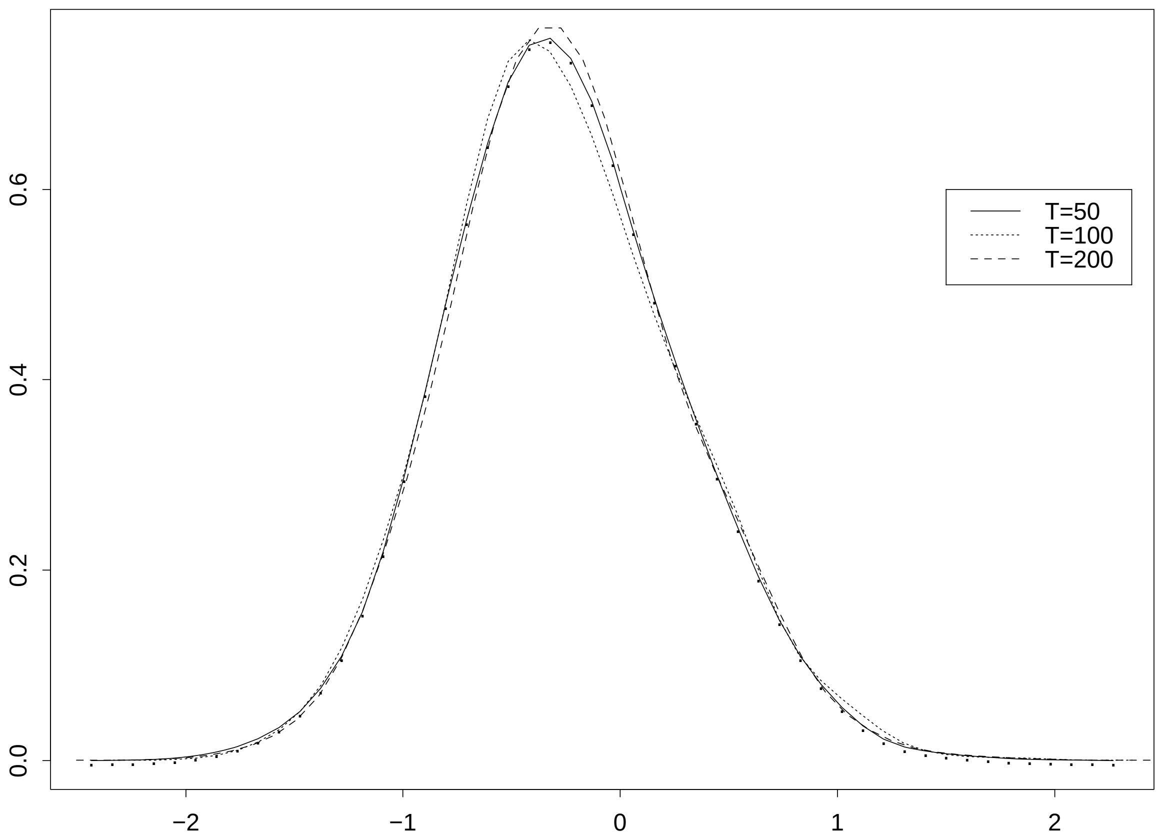


Figure 2b: Finite Sample Distribution of xi_T, Regression 2

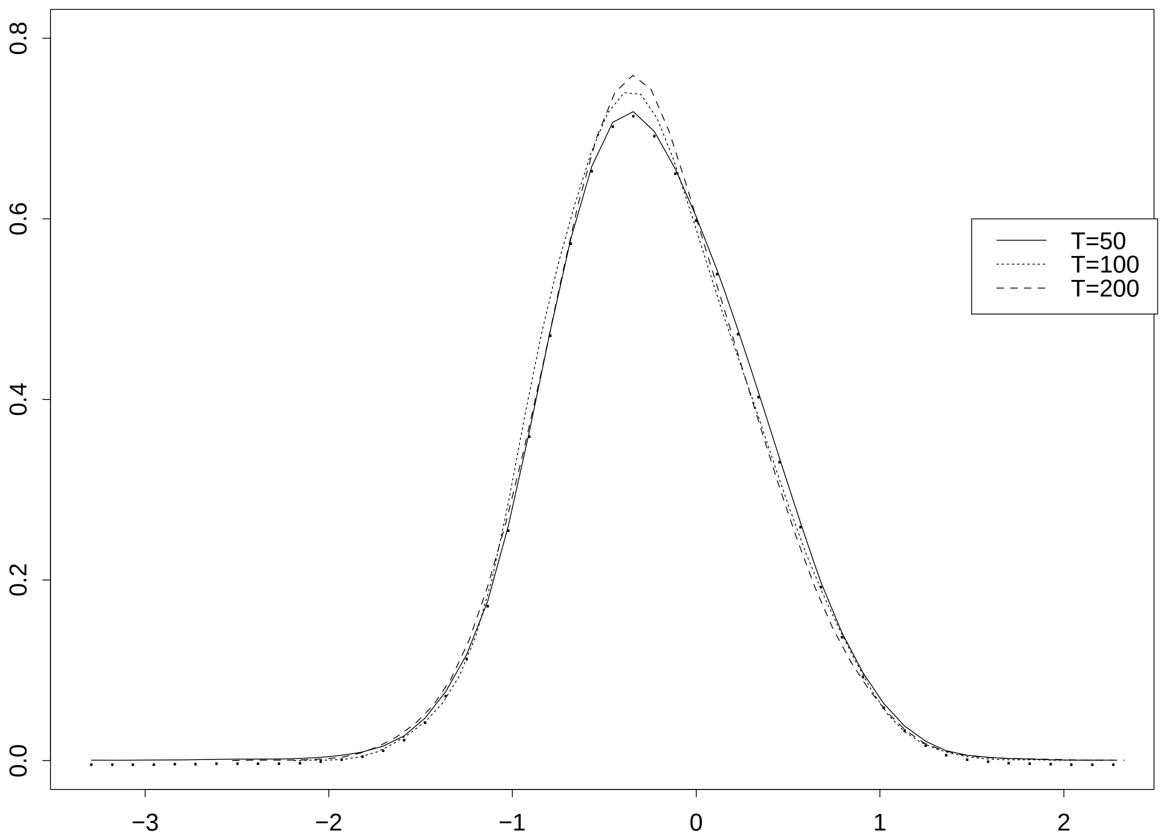


Figure 3a: Finite Sample Distribution of eta(alpha)_T, Reg.1

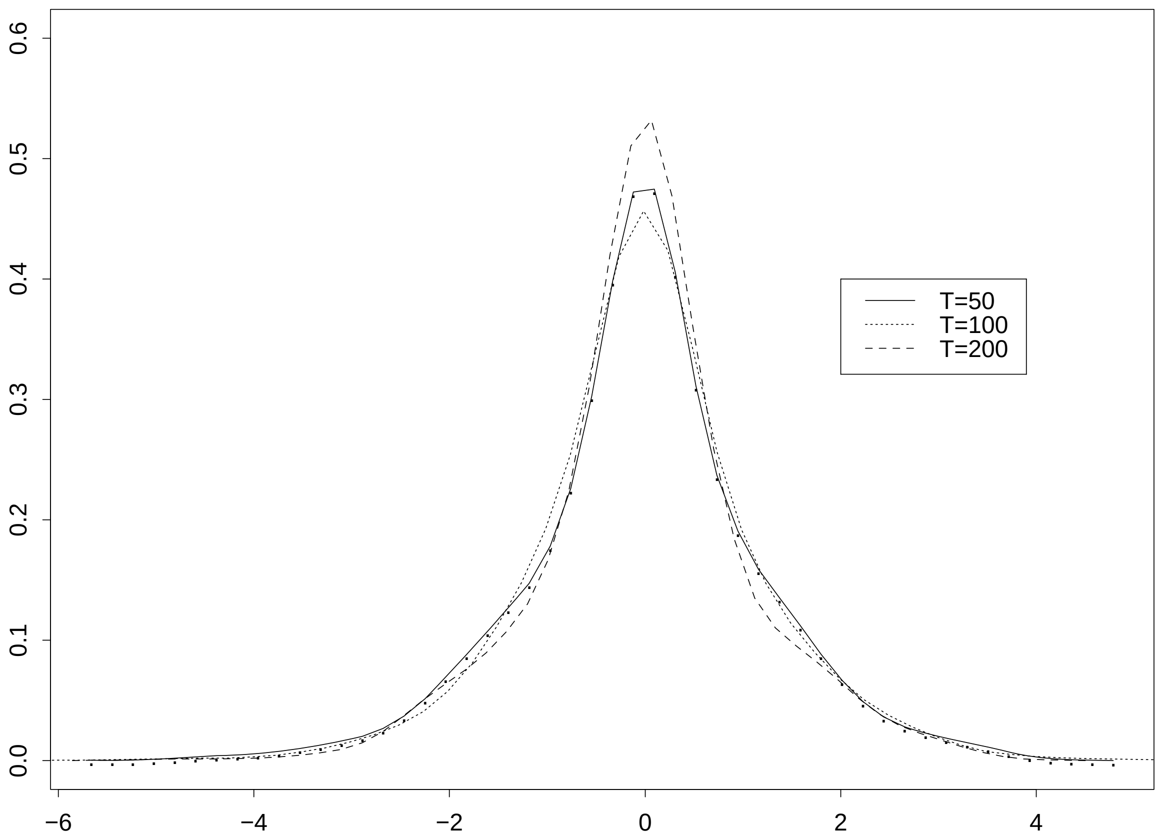


Figure 3b: Finite Sample Distribution of eta(alpha)_T, Reg.2

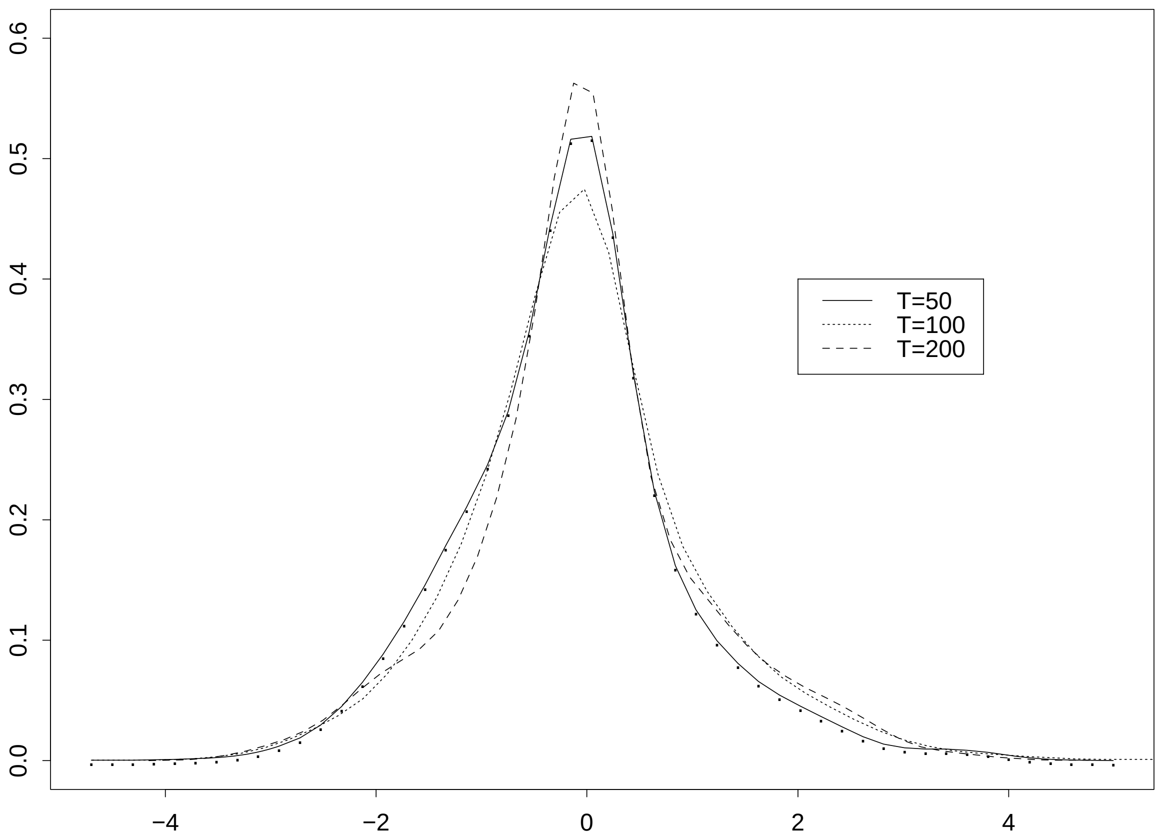


Figure 4a: Finite Sample Distribution of eta(beta)_T, Reg.1

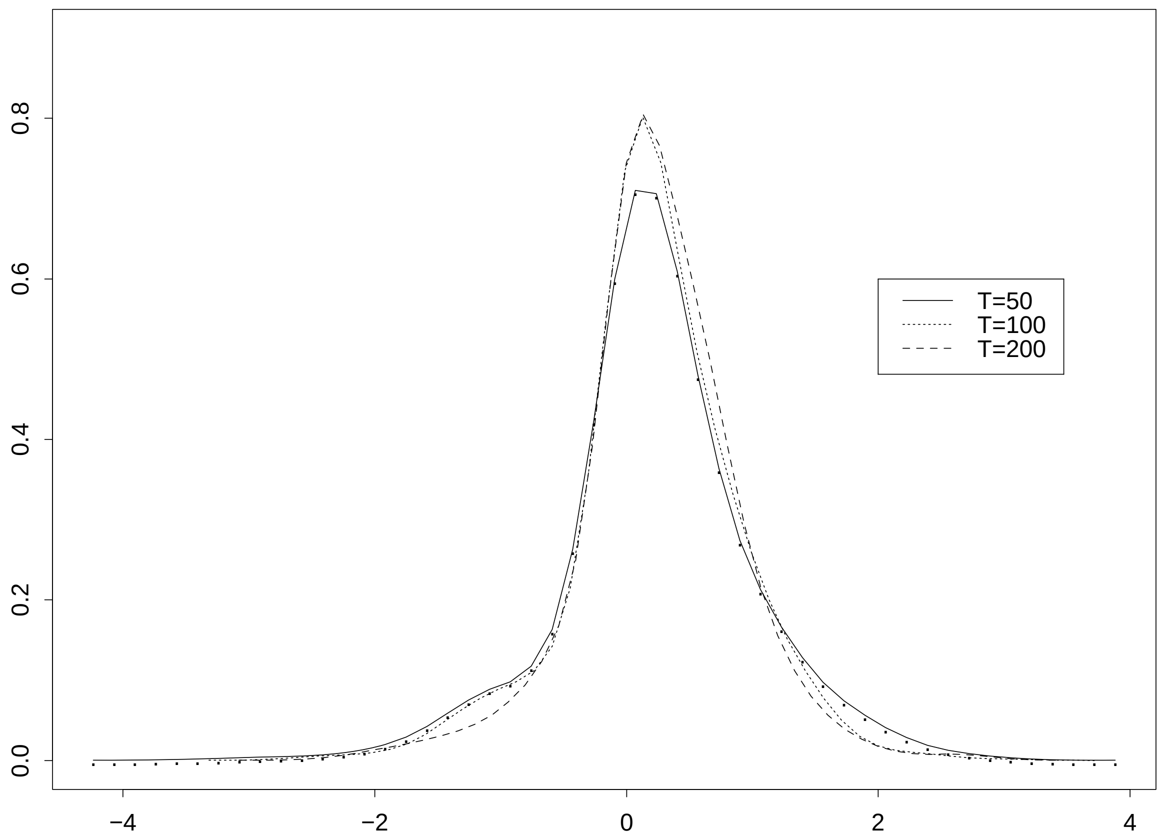
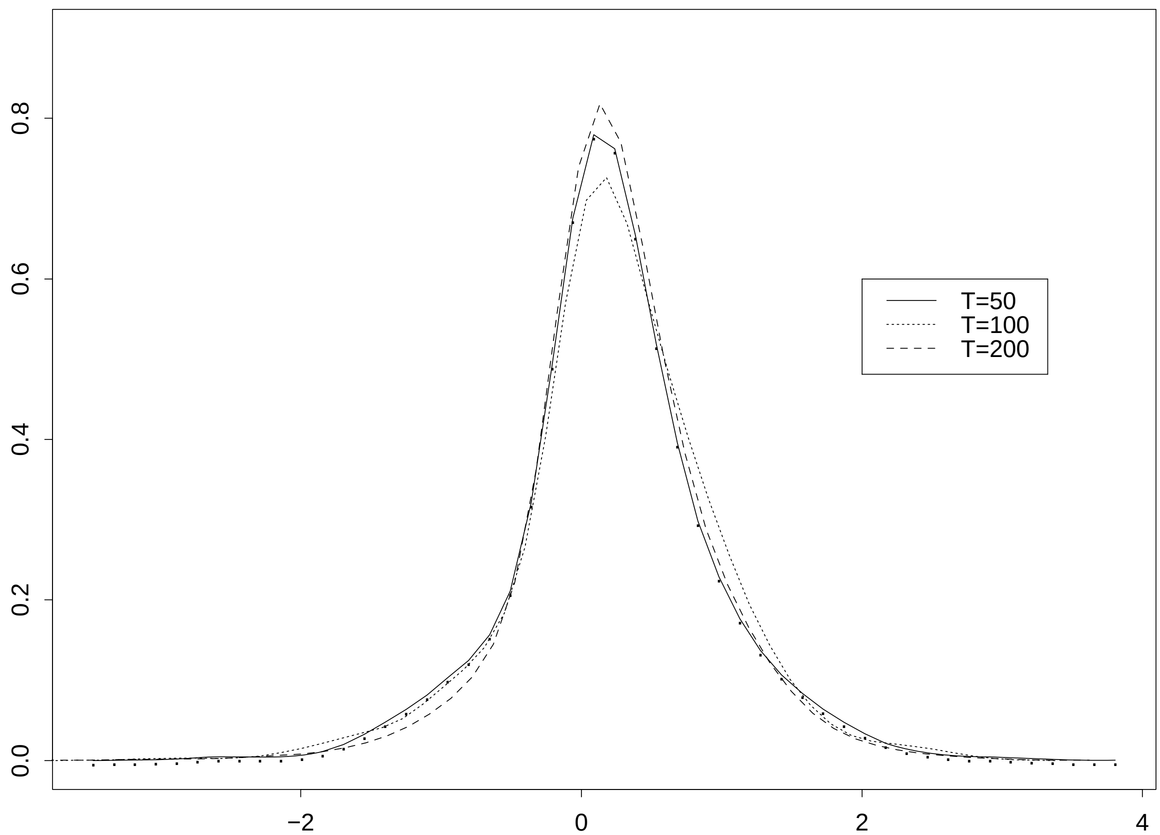


Figure 4b: Finite Sample Distribution of eta(beta)_T, Reg.2



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., and P., Bernstein (2002): "What Risk Premium is Normal?", Financial Analysts Journal, 58, 64-85.

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

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

[9] Bilodeau, M., and D., Brenner (1999): "Theory of Multivariate Statistics", Springer-Verlag.

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

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

[12] Boudoukh, J., Richardson, M., and 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., and L., Zhang (2006): "Equity Market Volatility and Expected Risk Premium", Federal Reserve Bank of St. Louis.

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

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

[16] Engle, R., Granger, C., and D., Kraft (1984): " Combining Competing Forecats of Inflation Using A Bivariate ARCH Model", Journal of Economic Dynamics and Control, 8, 151-165.

[17] Engle, R., and K., Kroner (1995): "Multivariate Simultaneous Generalized ARCH", Econometric Theory, 11, 122-150.

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

[19] Gourieroux, C. (2006): "Positivity Condition for a Bivariate Autoregressive Volatility Specification", CREST DP.

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

[21] Gourieroux, C., Holly, A., and 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.

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

[23] Gourieroux, C., Jasiak, J., and R., Sufana (2004): "The Wishart Autoregressive Process of Multivariate Stochastic Volatility",

[24] Gourieroux, C., and A., Monfort (1989): "Statistique et Methodes Econometriques", Vol 2, Paris, Economica. English translation (1995): "Statistics and Econometric Models", Vol 2, Cambridge Univ. Press, Cambridge UK.

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

[26] Guggenberger, P. Hahn, J., and K., Kim (2006): "Specification Testing Under Moment Inequalities", UCLA DP.

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

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

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

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

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

[32] Moon, R., and F., Schorfheide (2005): "Empirical Likelihood Estimation with Inequality Moment Restrictions", University of Pensylvania D.P.

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

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

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

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

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