Finite-sample quantiles of the Jarque-Bera test

Steve Lawford*
Department of Economics and Finance, Brunel University

First draft: February 2004.

Abstract

The finite-sample null distribution of the Jarque-Bera Lagrange multiplier test for normality differs considerably from the asymptotic $\chi^2(2)$. However, asymptotic critical values are commonly used in applied work, even for relatively small sample sizes. Here, we develop very accurate response surface approximations for the 10% and 5% critical values of the test, which enable correct practical implementation.

Keywords: Finite-sample critical values; Jarque-Bera test; Monte Carlo simulation; Normality; Response surfaces.
JEL classification: C12; C15; C52

*Corresponding author: Steve Lawford, Department of Economics and Finance, Brunel University, Uxbridge, UB8 3PH, UK; steve_lawford@yahoo.co.uk; fax: +44 1895 203 384.

1 Introduction

The Jarque-Bera (1980,1987) Lagrange multiplier test is perhaps the most commonly used procedure for testing whether a univariate sample of T datapoints, or estimated regression residuals, are drawn from a normal distribution. It is a joint test of the null hypothesis (of normality) that sample skewness equals 0 and sample kurtosis equals 3, and the null is rejected when the statistic

$$\mathrm{LM} = \frac{T}{6}\left\{\left(b_1^{1/2}\right)^2 + \frac{(b_2-3)^2}{4}\right\} \overset{a}{\sim} \chi^2(2)$$

exceeds some critical value, which is usually taken from the asymptotic $\chi^2(2)$ distribution. The standardized third and fourth moments are given by $b_1^{1/2} = m_3/m_2^{3/2}$ and $b_2 = m_4/m_2^2$ respectively, and m_i is the ith central moment of the sample. It has been noted that the small-sample tail quantiles of the LM statistic are quite different from their asymptotic counterparts; e.g. Deb and Sefton (1996, Table 1) and Urzúa (1996, Table 1). The use of asymptotic critical values given even fairly large samples will distort the actual size of the test, and may lead to incorrect decisions in applied work.

Deb and Sefton (1996) compute 14 very accurate empirical 10% and 5% significance points of LM in the interval $T \in [20, 800]$, and show that their use gives an almost correctly-sized test using regression residuals, when various data generating processes are chosen for the regressors. However, practical implementation using their critical values requires new simulations for sample sizes that are not tabulated. We address this problem, and develop highly accurate response surface[1] approximations to the 10% and 5% finite-sample critical values of LM, that are generally correct to ± 0.01, and may be used for $T \geq 5$.

2 Finite-sample critical values

Using Monte Carlo simulation, we generate 1000000 realizations of LM under the null of normality, for each sample size T in the set

$$T \in \{5, 6, \ldots, 25, 30, \ldots, 100, 125, \ldots, 1000\}, \ \alpha \in \{0.90, 0.95\}, \qquad (1)$$

and calculate the 10% and 5% critical values as the $\alpha\,(1000000)$th largest values of LM. This procedure gives 72 datapoints for each α, which are rather

[1]Response surfaces are numerical-analytical approximations, that have been widely applied in econometrics; e.g. Ericsson (1991), Cheung and Lai (1995), MacKinnon (1994), MacKinnon, Haug and Michelis (1999), and Ericsson and MacKinnon (2002).

more accurate than those previously available. We generate standard normal pseudorandom numbers similarly to MacKinnon's (1994, p. 170) long-period algorithm. The sample sizes that we have chosen are representative of those that are commonly used in applied work. Our design focuses on $T \leq 25$, since the actual critical values vary widely across this range, and specification of good response surfaces requires this information. All simulations were performed on a Pentium 4 machine, with a 2GHz processor and 256MB of RAM, running GAUSS under Microsoft Windows XP.

We regressed the Monte Carlo estimates of $100\alpha\%$ quantiles on various functions of sample size, constructed so that $T \to \infty$ gives the (known) asymptotic quantiles. Following much experimentation, and motivated by the quantile approximations developed by MacKinnon, Haug, and Michelis (1999, p. 569) in the context of Johansen-type tests for cointegration, we chose to fit the following quantile response surface:

$$q^{\alpha}(T_i) = q^{\alpha}_{\infty} + \sum_{k=1}^{9} \beta_k T_i^{-k} + u_i. \tag{2}$$

The dependent variable $q^{\alpha}(T_i)$ is the simulated finite-sample $100\alpha\%$ quantile with sample size T_i, which takes values from (1); q^{α}_{∞} is the asymptotic $100\alpha\%$ quantile from the $\chi^2(2)$ distribution, which was computed in GAUSS as $q^{\alpha}_{\infty} = \arg\min \{\text{cdfchic}(q, 2) - \alpha + 1\}^2$; u_i is an error term.

We denote the estimated response surface by $\widehat{q}^{\alpha}$, and estimated coefficients are reported in Table 1. Selection criteria included small residual variance, parsimony, and satisfactory diagnostic performance. The dependencies of the 10% and 5% critical values on sample size are presented in Figures 1–3, which plot response surfaces $\widehat{q}^{\alpha}$ against T. The response surface fits are very good, and generally agree with the simulated quantiles to roughly ± 0.01, across the entire parameter space (1). For instance, estimated critical values are 2.75 (10%) and 4.41 (5%) for a sample size of 30, and 3.48 (10%) and 5.28 (5%) for a sample size of 75. Simulated critical values are 2.74 (10%) and 4.41 (5%) for a sample size of 30, and 3.49 (10%) and 5.27 (5%) for a sample size of 75. We considered more parsimonious approximations, with fewer inverse powers of T_i, although these failed to yield an improved fit over (2). Clearly, $\widehat{q}^{0.90}$ and $\widehat{q}^{0.95}$ break-down for $T \leq 4$, although this is unlikely to be a problem in applied work.

3 Concluding comments

We have developed very accurate response surface approximations to the 10% and 5% critical values of the Jarque-Bera test for normality, that are computationally simple, and may be used to give an almost correctly-sized test in empirical work.

Acknowledgements This paper was typed in Scientific Word and numerical results were derived using GAUSS and E–Views.

References

1. Cheung, Y. -W., Lai, K. S., 1995. Lag order and critical values of the augmented Dickey-Fuller test, Journal of Business and Economic Statistics 13, 277-280.

2. Deb, P., Sefton, M., 1996. The distribution of a Lagrange multiplier test of normality, Economics Letters 51, 123-130.

3. Ericsson, N. R., 1991. Monte Carlo methodology and the finite sample properties of instrumental variables statistics for testing nested and non-nested hypotheses, Econometrica 59, 1249-1277.

4. Ericsson, N. R., MacKinnon, J. G., 2002. Distributions of error correction tests for cointegration, Econometrics Journal 5, 285-318.

5. Jarque, C. M., Bera, A. K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals, Economics Letters 6, 255-259.

6. Jarque, C. M., Bera, A. K., 1987. A test for normality of observations and regression residuals, International Statistical Review 55, 163-172.

7. MacKinnon, J. G., 1994. Approximate asymptotic distribution functions for unit-root and cointegration tests, Journal of Business and Economic Statistics 12, 167-176.

8. MacKinnon, J. G., Haug, A. A., Michelis, L., 1999. Numerical distribution functions of likelihood ratio tests for cointegration, Journal of Applied Econometrics 14, 563-577.

9. Urzúa, C. M., 1996. On the correct use of omnibus tests for normality, Economics Letters 53, 247-251 (corrigendum, 1997, 54: 301).

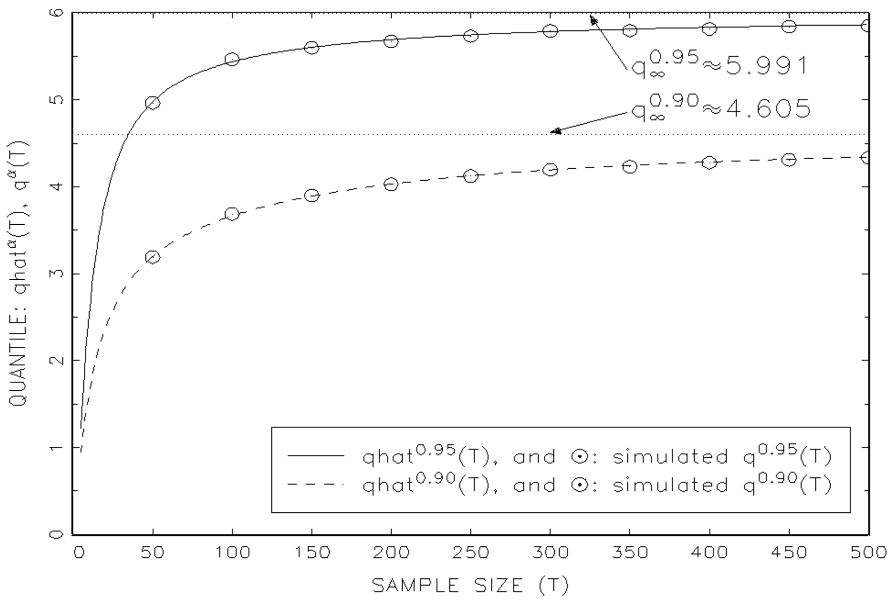

Fig. 1. Quantile response surfaces $\widehat{q}^{0.95}$ and $\widehat{q}^{0.90}$, for $T \in [5, 500]$.

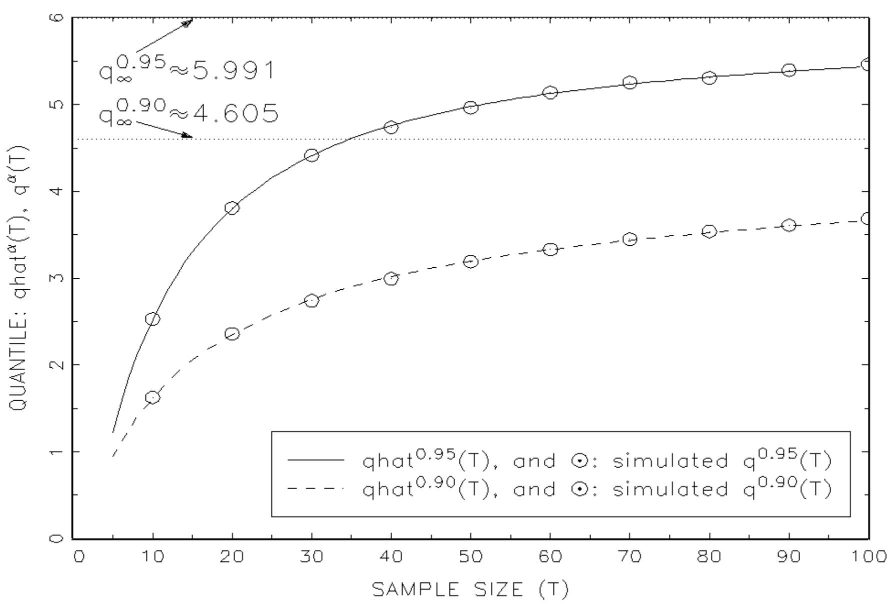

Fig. 2. Quantile response surfaces $\widehat{q}^{0.95}$ and $\widehat{q}^{0.90}$, for $T \in [5, 100]$.

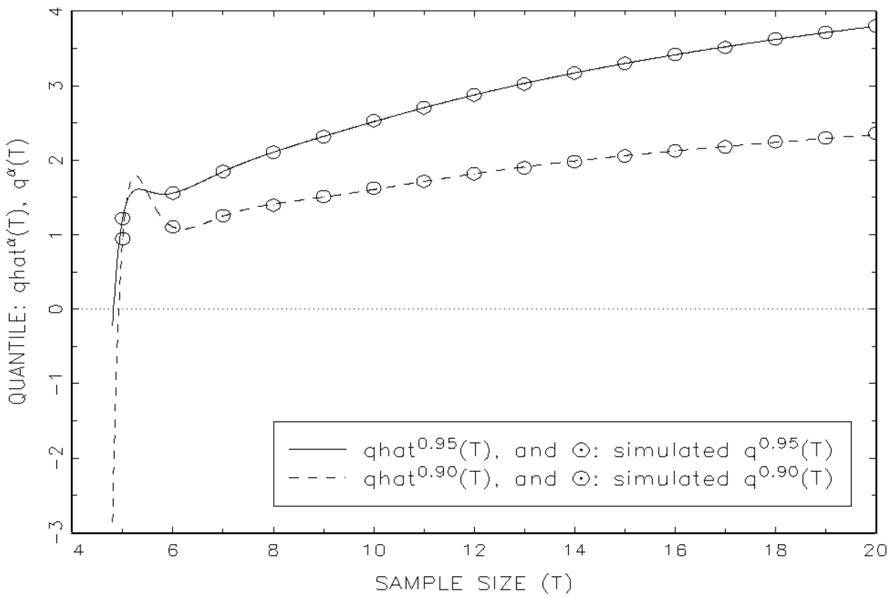

Fig. 3. Quantile response surfaces $\widehat{q}^{0.95}$ and $\widehat{q}^{0.90}$, for $T \in [5, 20]$.

Table 1[2]: Estimated quantile response surfaces $\widehat{q}^{\alpha}$		
	$\alpha = 0.90$	$\alpha = 0.95$
q_{∞}^{α}	4.605049	5.9913104
$\widehat{\beta}_1$	−145.1816602	−67.00449919
$\widehat{\beta}_2$	7 286.233799	1 719.108744
$\widehat{\beta}_3$	−275 153.9753	−74 443.10488
$\widehat{\beta}_4$	6 437 304.253	1 962 801.944
$\widehat{\beta}_5$	−92 456 006.82	−30 095 541.45
$\widehat{\beta}_6$	814 503 598.1	275 285 058.6
$\widehat{\beta}_7$	−4 276 230 401	−1 479 198 621
$\widehat{\beta}_8$	12 243 649 840	4 299 485 882
$\widehat{\beta}_9$	−14 677 406 860	−5 206 421 393
$\overline{R}^2$	0.9999	0.9999
RSS	0.006423	0.009567
Mean $\lvert\widehat{u}\rvert$	0.00782	0.00899
Max $\lvert\widehat{u}\rvert$	0.0230	0.0344

[2] The response surfaces (2) were estimated in E–Views. All estimated coefficients were significant at the 1% level. $\overline{R}^2$ is the degrees-of-freedom adjusted coefficient of determination. RSS is the residual sum of squares. Mean $|\widehat{u}|$ is the mean absolute error of the response surface approximations against the simulated critical values, and Max $|\widehat{u}|$ is the maximum absolute error.