Precise finite-sample quantiles of the Jarque-Bera adjusted Lagrange multiplier test

Wuertz, Diethelm and Katzgraber, Helmut

Swiss Federal Institute of Technology, Zurich

11 December 2009

Precise finite-sample quantiles of the Jarque-Bera adjusted Lagrange multiplier test

Diethelm Würtz and Helmut G. Katzgraber

Swiss Federal Institute of Technology
Institute for Theoretical Physics
ETH Hönggerberg, CH-8093 Zürich

wuertz@itp.phys.ethz.ch, katzgraber@phys.ethz.ch

Abstract:

It is well known that the finite-sample null distribution of the Jarque-Bera Lagrange Multiplier *(LM)* test for normality and its adjusted version *(ALM)* introduced by Urzua differ considerably from their asymptotic $\chi^2(2)$ limit. Here, we present results from Monte Carlo simulations using 10^7 replications which yield very precise numbers for the *LM* and *ALM* statistic over a wide range of critical values and sample sizes. Depending on the sample size and values of the statistic we get p values which significantly deviate from numbers previously published and used in hypothesis tests in many statistical software packages. The p values listed in this short Letter enable for the first time a precise implementation of the Jarque-Bera *LM* and *ALM* tests for finite samples.

1 Introduction

The Jarque-Bera (1980, 1987) Lagrange multiplier test is likely the most widely used procedure for testing normality of economic time series returns. The algorithm provides a joint test of the null hypothesis of normality in that the sample skewness b_1 equals zero and the sample kurtosis b_2 equals three. The null is rejected when the Lagrange multiplier statistic

$$LM = N\left(\frac{(b_1^{1/2})^2}{6} + \frac{(b_2-3)^2}{24}\right) \tag{1}$$

exceeds some critical value, which is taken in the asymptotic limit from the $\chi^2(2)$ distribution. N is the sample size, $b_1^{1/2} = m_3/m_2^{3/2}$, $b_2 = m_4/m_2^2$ where m_i is the i-th central moment of the observations $m_i = \Sigma(x_j - \overline{x})^i/N$, and $\overline{x}$ the sample mean.

Urzua (1996) modified the Jarque-Bera test replacing the asymptotic means and variances by their exact finite-sample values yielding

$$ALM = N\left(\frac{(b_1^{1/2})^2}{c_1} + \frac{(b_2-c_2)^2}{c_3}\right) \ . \tag{2}$$

Here the parameters $c_{1,2,3}$ are given by the expectation value and variances of the skewness and kurtosis

$$c_1 = var(b_1^{1/2}) = \frac{6(N-2)}{(N+1)(N+3)} \ ,$$

$$c_2 = E(b_2) = \frac{3(N-1)}{(N+1)} \ ,$$

$$c_3 = var(b_2) = \frac{24N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)} \ .$$

Note, that the ALM has the same asymptotic distribution as the LM statistic.

The work of Urzua (1996) as well as the work by Deb and Sefton (1996) already warn about the incorrect use of the Jarque-Bera test in the case of small- and medium-sized samples. The authors performed Monte Carlo simulations and tabulated significance points for 5% and 10%, on a series of sample sizes ranging between 10 and 800. Deb and Sefton used 600'000 replications in their Monte Carlo simulations and Urzua used 10'000 replications and added results for the 1%, 15% and 20% significance points. Very recently Lawford (2004) developed an accurate response surface approximation for the 5% and 10% critical values of the Jarque-Bera test based on Monte Carlo simulations using 1 Million replications. The tables for the LM and ALM statistic values presented in these papers are restricted usually to a small set of parameters and the precision is in most cases limited to two digits. Furthermore, for small N we observe significant differences in comparison to previously published values. For some parameter settings the differences are so large, that this may result in inaccurate hypothesis tests or evenmore this may lead to situations with wrong decisions.

In this Letter we present tables with very precise values for both, the LM and ALM statistic. Since the slow convergence of the Monte Carlo simulation is well known we extend the simulations to 10 Million replications and enhance the mesh of p-values and sample sizes considerably.

The results have been used to implement R functions for the finite sample Jarque-Bera test and the distribution itself, using either the LM or ALM statistic. R (2004) is a powerful and widely used GPL-licensed statistical software environment based on the S language. In this sense our functions can also be called from the commercial S-Plus software package. The R functions are part of the Rmetrics software project, *www.rmetrics.org*. The software is GPL licensed and can be downloaded from the CRAN Server *www.r-project.org*.

2 Monte Carlo Simulation

We performed Monte Carlo simulations of the LM and ALM statistic using 10^7 replications. The results are summarized in Table 1 for both the LM and ALM statistic.

LM: p\N	10	20	35	50	75	100	150	200	300	500	800	1000	1600	2400	10000
0.01%	15.345	46.996	66.612	71.734	69.910	68.032	60.632	54.736	47.572	38.847	33.247	31.213	26.956	24.249	19.940
0.05%	12.444	31.159	40.759	43.256	41.909	40.430	37.229	34.330	30.561	26.270	23.045	21.979	19.760	18.366	16.052
0.10%	10.995	24.970	31.969	33.753	32.738	31.840	29.547	27.551	24.830	21.812	19.521	18.736	17.150	16.083	14.397
0.50%	7.3004	13.471	16.414	17.281	17.305	16.959	16.257	15.638	14.669	13.583	12.726	12.366	11.762	11.384	10.792
1.00%	**5.7029**	**9.7182**	**11.736**	**12.392**	**12.586**	**12.491**	**12.185**	**11.882**	**11.3580**	**10.778**	**10.299**	**10.117**	**9.8095**	**9.6084**	**9.3128**
5.00%	**2.5247**	**3.7954**	**4.5929**	**4.9757**	**5.2777**	**5.4300**	**5.5984**	**5.6758**	**5.7732**	**5.8551**	**5.9103**	**5.9242**	**5.9569**	**5.9671**	**5.9857**
10.00%	**1.6232**	**2.3470**	**2.8814**	**3.1834**	**3.4862**	**3.6734**	**3.9041**	**4.0327**	**4.1891**	**4.3317**	**4.4274**	**4.4568**	**4.5132**	**4.5424**	**4.5888**
15.00%	1.2826	1.8230	2.2533	2.5094	2.7713	2.9390	3.1416	3.2580	3.4003	3.5312	3.6198	3.6507	3.7016	3.7309	3.7778
20.00%	1.1236	1.5623	1.9162	2.1278	2.3463	2.4865	2.6558	2.7559	2.8764	2.9882	3.0645	3.0909	3.1360	3.1611	3.2036
30.00%	0.9389	1.2516	1.4997	1.6466	1.7975	1.8944	2.0112	2.0807	2.1639	2.2427	2.2962	2.3153	2.3460	2.3650	2.3968
40.00%	0.8077	1.0360	1.2115	1.3128	1.4165	1.4828	1.5619	1.6087	1.6649	1.7175	1.7547	1.7679	1.7889	1.8024	1.8248
50.00%	0.6950	0.8574	0.9771	1.0447	1.1126	1.1563	1.2076	1.2385	1.2752	1.3101	1.3338	1.3420	1.3568	1.3655	1.3808
60.00%	0.5885	0.6948	0.7699	0.8114	0.8529	0.8800	0.9105	0.9292	0.9518	0.9732	0.9882	0.9931	1.0024	1.0085	1.0181
70.00%	0.4801	0.5378	0.5769	0.5985	0.6202	0.6348	0.6508	0.6610	0.6730	0.6851	0.6940	0.6965	0.7018	0.7056	0.7108
80.00%	0.3618	0.3777	0.3896	0.3969	0.4046	0.4105	0.4168	0.4213	0.4267	0.4325	0.4368	0.4376	0.4402	0.4421	0.4451
85.00%	0.2950	0.2938	0.2958	0.2982	0.3010	0.3044	0.3071	0.3096	0.3130	0.3163	0.3189	0.3194	0.3209	0.3221	0.3245
90.00%	0.2192	0.2047	0.2002	0.1997	0.1997	0.2006	0.2016	0.2024	0.2040	0.2060	0.2071	0.2074	0.2081	0.2089	0.2106
95.00%	0.1272	0.1084	0.1022	0.1005	0.0995	0.0996	0.0992	0.0995	0.1000	0.1005	0.1010	0.1012	0.1013	0.1019	0.1024
99.00%	0.0304	0.0230	0.0208	0.0203	0.0198	0.0197	0.0196	0.0196	0.0197	0.0198	0.0197	0.0199	0.0020	0.0020	0.0020
99.50%	0.0156	0.0116	0.0104	0.0101	0.0099	0.0098	0.0098	0.0098	0.0099	0.0099	0.0098	0.0099	0.0099	0.0099	0.0100
99.90%	0.0032	0.0023	0.0021	0.0020	0.0020	0.0019	0.0019	0.0019	0.0020	0.0019	0.0020	0.0020	0.0020	0.0020	0.0020
99.95%	0.0016	0.0012	0.0010	0.0010	0.0010	0.0009	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
99.99%	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002

ALM: p\N	10	20	35	50	75	100	150	200	300	500	800	1000	1600	2400	10000
0.01%	51.600	91.217	99.883	96.158	85.696	79.523	67.685	59.600	50.530	40.472	34.231	31.993	27.454	24.575	20.040
0.05%	41.502	60.508	60.927	58.013	51.413	47.444	41.713	37.524	32.579	27.431	23.751	22.545	20.128	18.621	16.114
0.10%	36.538	48.399	47.780	45.318	40.266	37.414	33.106	30.158	26.498	22.785	20.147	19.244	17.474	16.314	14.457
0.50%	23.831	25.963	24.569	23.229	21.334	19.986	18.285	17.156	15.689	14.211	13.129	12.694	11.971	11.525	10.827
1.00%	**18.374**	**18.643**	**17.540**	**16.659**	**15.506**	**14.719**	**13.707**	**13.042**	**12.149**	**11.271**	**10.616**	**10.372**	**9.9667**	**9.7158**	**9.3386**
5.00%	**7.4161**	**6.9317**	**6.6788**	**6.5533**	**6.4144**	**6.3192**	**6.2182**	**6.1493**	**6.0925**	**6.0497**	**6.0309**	**6.0218**	**6.0182**	**6.0077**	**5.9961**
10.00%	**4.1769**	**3.9657**	**3.9612**	**3.9977**	**4.0664**	**4.1256**	**4.2180**	**4.2718**	**4.3547**	**4.4336**	**4.4923**	**4.5095**	**4.5462**	**4.5650**	**4.5941**
15.00%	2.8110	2.7736	2.8895	2.9935	3.1215	3.2150	3.3356	3.4086	3.5045	3.5952	3.6611	3.6833	3.7228	3.7450	3.7812
20.00%	2.1830	2.2164	2.3547	2.4616	2.5881	2.6767	2.7895	2.8582	2.9462	3.0321	3.0923	3.1132	3.1502	3.1706	3.2058
30.00%	1.6376	1.6569	1.7585	1.8388	1.9325	1.9986	2.0826	2.1350	2.2003	2.2646	2.3098	2.3262	2.3528	2.3694	2.3980
40.00%	1.3166	1.3130	1.3785	1.432	1.4975	1.5435	1.6021	1.6382	1.6840	1.7287	1.7613	1.7731	1.7917	1.8042	1.8251
50.00%	1.0658	1.0460	1.0844	1.1183	1.1604	1.1904	1.2290	1.2536	1.2846	1.3150	1.3367	1.3441	1.3579	1.3663	1.3808
60.00%	0.8464	0.8165	0.8344	0.8532	0.8781	0.8971	0.9200	0.9359	0.9554	0.9746	0.9888	0.9936	1.0027	1.0087	1.0179
70.00%	0.6406	0.6065	0.6100	0.6183	0.6309	0.6416	0.6538	0.6625	0.6735	0.6848	0.6937	0.6963	0.7014	0.7054	0.7108
80.00%	0.4376	0.4056	0.4005	0.4022	0.4064	0.4111	0.4163	0.4203	0.4257	0.4316	0.4359	0.4369	0.4397	0.4420	0.4451
85.00%	0.3344	0.3061	0.2994	0.2991	0.3005	0.3029	0.3058	0.3082	0.3117	0.3154	0.3181	0.3187	0.3204	0.3219	0.3243
90.00%	0.2284	0.2060	0.1992	0.1982	0.1979	0.1991	0.2000	0.2012	0.2029	0.2051	0.2065	0.2071	0.2078	0.2087	0.2105
95.00%	0.1177	0.1044	0.0996	0.0986	0.0980	0.0981	0.0982	0.0985	0.0993	0.1001	0.1007	0.1009	0.1013	0.1017	0.1023
99.00%	0.0242	0.0213	0.0199	0.0197	0.0194	0.0194	0.0193	0.0195	0.0195	0.0196	0.0196	0.0198	0.0199	0.0199	0.0200
99.50%	0.0122	0.0106	0.0100	0.0098	0.0097	0.0097	0.0097	0.0097	0.0097	0.0098	0.0098	0.0099	0.0099	0.0099	0.0100
99.90%	0.0025	0.0021	0.0020	0.0020	0.0020	0.0019	0.0019	0.0019	0.0019	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020
99.95%	0.0013	0.0011	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
99.99%	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002

Table 1: Top: Significance points for the finite sample Jarque-Bera test. Bottom: Same values for the adjusted Jarque-Bera Test. The numbers are based on Monte Carlo simulations using 10^7 replications. Note, that the p values are listed in reverse order as $1 - p$. The three major levels, 1%, 5% and 10%, are written in bold face.

Figure 1 illustrates the results in a graph. The simulated p values and the deviations from the asymptotic $\chi^2(2)$ limit are shown. The curves belong to the same values of sample sizes N as listed in table 1.

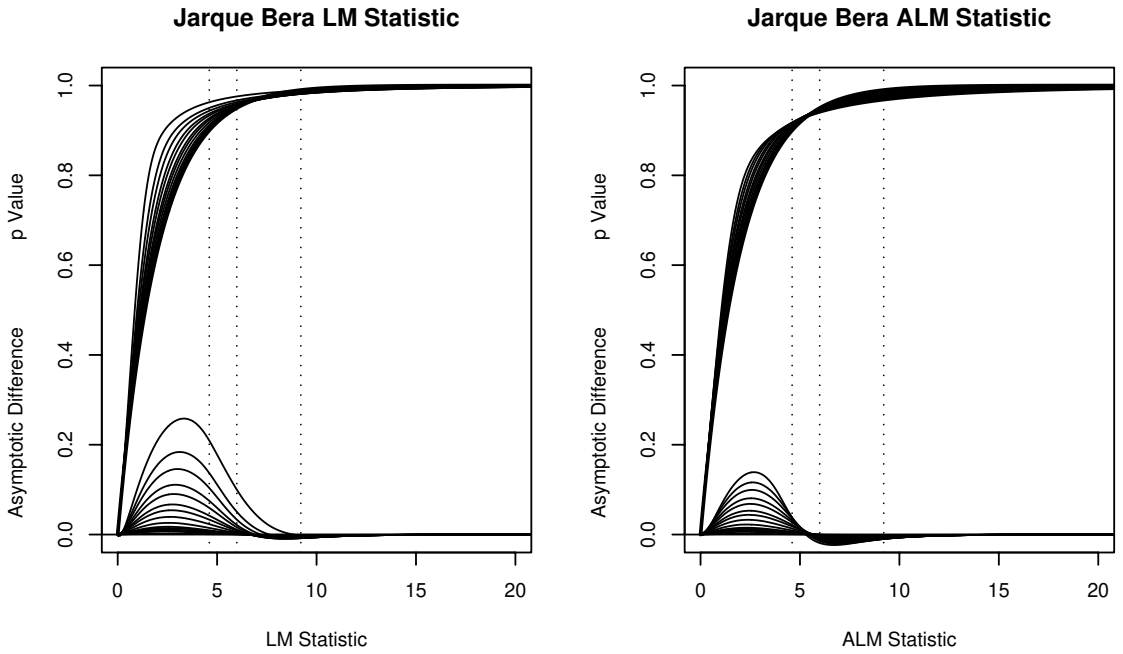

Figure 1: LM (left) and ALM (right) finite sample p values and their differences with respect to the asymptotic limit. The upper bundle of curves shows the p values. The lower bundle of curves measures the difference $p_N - p_\infty$ to the asymptotic limit. The graph clearly demonstrates that the adjusted Jarqua-Bera test outperforms the original version of the test. The three dotted vertical lines mark the 1% (99%), 5% (95%) and 10% (95%) levels in the asymptotic limit, respectively.

3 Response Surface and Hypothesis Test

To compute the LM and ALM statistic for a wide range of quantiles and sample sizes one usually approximates the response surface for a fixed value of p as a series in powers of 1 over N

$$q(p, N) = q(p, \infty) + \sum_{k=1}^{K} \beta_k N^{-k} \ . \tag{3}$$

Lawford (2004) has done this for the 5% and 10% quantile lines. He fitted his Monte Carlo data based on 1 Million replications for $K = 9$. The regression coefficients β are listed in the aforementioned paper. We have done fits over a wide range of p-values. The results are shown in figure 2 in comparison with those obtained by Lawford. Note that Lawford's fit becomes less reliable for small lengths where the convergence of the Monte Carlo simulation slows down.

Another approach would be an Edgeworth (1917) expansion of the distribution in $1/N$. Unfortunately, we found out that the expansion converges extremely slow. So we applied "Curve Fitting", as suggested by Rothenberg (1984), to approximate the response surface. Simple linear interpolation, 2-dimensional splines or connectionist function approximators are only three possibilities from many others. We followed the first approach fitting on logarithmic scales. The results are shown in Figure 3 for both the traditional Jarque-Bera test as well as its adjusted version.

We have implemented the Jarque-Bera test for finite samples into S functions using the statistical software packages R and SPlus, but it can be done very easily in any other software environment like Matlab, Eviews, or SAS among others. The underlying simulations with 10^7 replications were done with a separate C program using a multiplicative lagged Fibonacci random number generator with a lag of size 1279. The software allows to compute the distribution function and the quantile function for finite samples and the asymptotic limit either for the LM or ALM test version. These functions are used to derive the p values by the hypothesis test function.

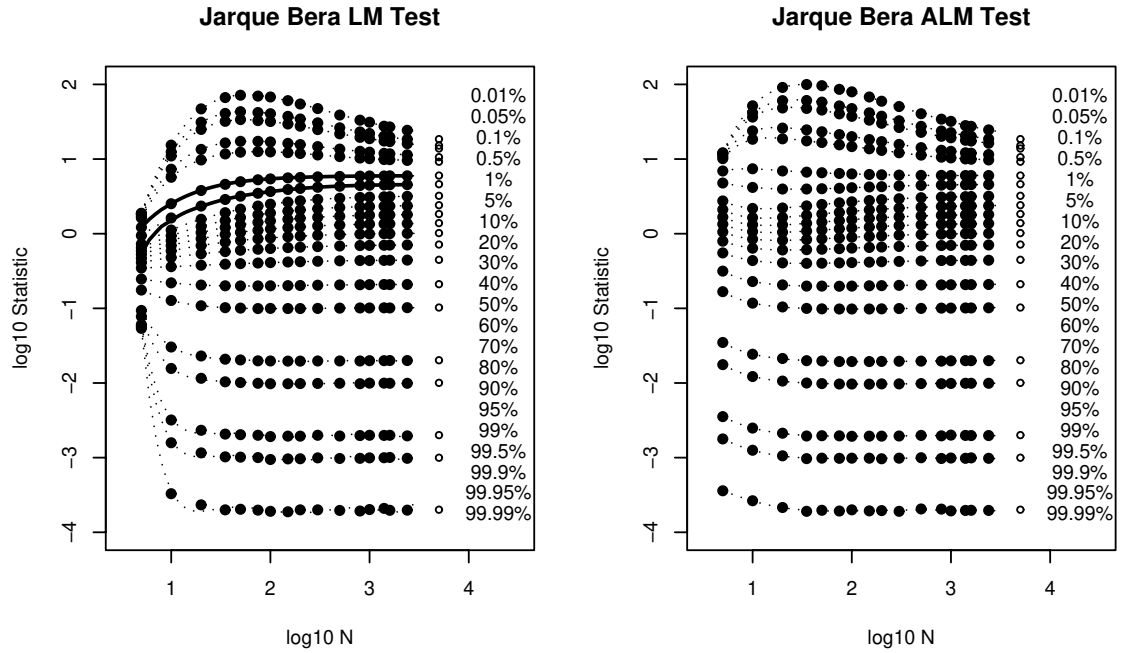

Figure 2: The figures show the LM (left) and ALM (right) statistic for a wide range of p values as a function of sample sizes. The dots show the results from the Monte Carlo simulations using 10^7 *replications together with the asymptotic limit (marked by the open circles). The dotted lines are fitted series expansions of order* $K = 6$ *in* $1/N$*. The two thick lines in the left LM graph display the results of Lawford for the 5% and 10% levels.*

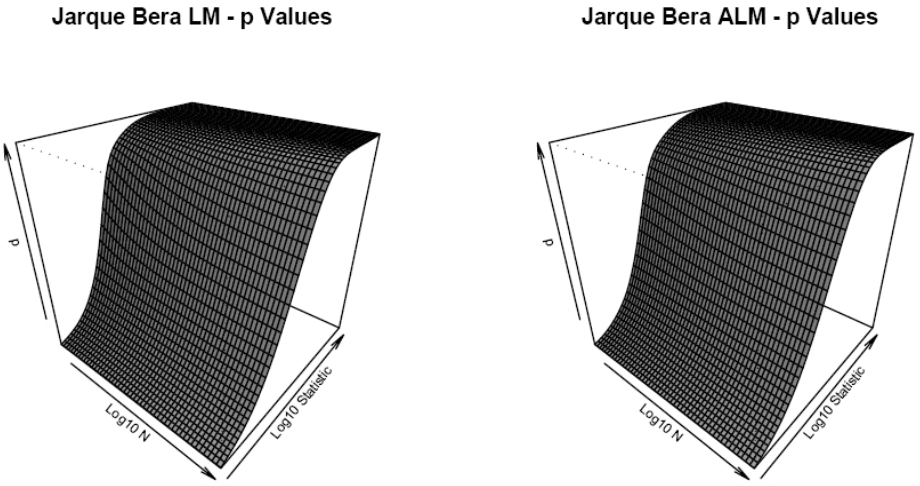

Figure 3: The figures show the LM (left) and ALM (right) surface of p values for a wide range of statistics (0.4 ... 100) and sample sizes (10 ... 10'000). Note, that the x- and y-axis are on logarithmic scales. The inputs consist of almost 2000 p-values ranging between 0.0001 *and* 0.9999*.*

4 Summary

This Letter tabulates precise p-values for the Jarque-Bera finite sample normality test. In addition to the original version of the Lagrange Multiplier test we have also computed finite sample p-values for its adjusted version formulated by Urzua (1996). In contrast to previous investigations the results were derived from a MC simulation with 10^7 replications. To our knowledge this is one of the largest simulations ever done in statistics. The outcome of the simulation are very precise values for finite samples which we have tabulated and can now be used for an improved hypothesis testing.

References

Bowman K., Shenton L., *Omnibus test contours for departures from normality based on b_1 and b_2*, Biometrika 62, 1975, 243–250.

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

Edgeworth F.Y., *On the mathematical representation of statistical data*, Journal of the Royal Statistical Society 80, , 1917, 411–437.

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

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

Lawford S., *Finite-sample quantiles of the Jarque-Bera test*, Brunel University Preprint, 2004, 7 pages.

R Core Team, *R Manuals*, downloadable from: http://cran.r-project.org.

Rmetrics, *Teaching Financial Engineering and Computational Finance with R*, http://www.rmetrics.org.

Rothenberg T.J., *Approximating the distributions of econometric estimators and test statistics*, Handbook of Econometrics, Volume II, 1984, 881–935.

Urzua M., *On the correct use of omnibus tests for normality*, Economics Letters 53, 1996, 247–251.