A Robust Entropy-Based Test for Asymmetry*

Esfandiar Maasoumi
Department of Economics
Southern Methodist University
Dallas, TX 75275-0496, USA
maasoumi@mail.smu.edu

Jeff Racine
Department of Economics
Syracuse University
Syracuse, NY 13244-1020, USA
jracine@maxwell.syr.edu

May 29, 2003

Abstract

We consider a metric entropy capable of detecting deviations from symmetry. A consistent test statistic is constructed from an integrated normed difference between two density functions estimated using kernel methods. The null distribution (symmetry) is obtained by resampling from an artificially lengthened series constructed from a rotation of the original series about its mean (median, mode). Simulations demonstrate that the test has correct level and power in the direction of interesting alternatives, while applications to updated Nelson & Plosser (1982) data demonstrate its potential power gains relative to existing tests.

1 Overview

Testing for asymmetric behavior present in a series or in conditional predictions thereof has a rich history dating to the pioneering work by Crum (1923), Mitchell (1927), and Keynes (1936) who examined potential asymmetries present in a number of macroeconomic series, while interest continues through the present day as exemplified by the recent work of Timmermann & Perez-Quiros (2001) and Belaire-Franch & Peiro (2003). Some researchers subdivide asymmetry into categories such as *sharpness*, *steepness*,

*Racine would like to thank the Center for Policy Research at Syracuse University for their ongoing support.

and *deepness* (see McQueen & Thorley (1993)). Though such categorizations of asymmetry may be of interest in their own right, our focus will rest upon consistent tests of asymmetries of any sort.

2 Unconditional and Conditional Symmetry

Consider a stationary series $\{Y_t\}_{t=1}^{T}$. Let $\mu_y = E[Y_t]$, let $f(y)$ denote the density function of the random variable Y_t, let $\tilde{Y}_t = -Y_t + 2\mu_y$ denote a rotation of Y_t about its mean, and let $f(\tilde{y})$ denote the density function of the random variable $\tilde{Y}_t$. Note that if $\mu_y = 0$ then $\tilde{Y}_t = -Y_t$, though in general this will not be so.

We say a series is *symmetric about the mean* (median, mode) if $f(y) \equiv f(\tilde{y})$ almost surely. Tests for asymmetry about the mean therefore naturally involve testing the following null:

$$H_0 : f(y) = f(\tilde{y}) \text{ for all } y. \tag{1}$$

One questions why the mean has received particular attention when symmetry about the mode or median would seem a more natural characterization. One could of course clearly rotate a distribution around these measures of central tendency, and for what follows one simply would replace the mean with the appropriate statistic.

Tests for the presence of conditional asymmetry can be based upon standardized residuals from a regression model (see Belaire-Franch & Peiro (2003)). Let

$$Y_t = h(\Omega_t, \beta) + \sigma(\Omega, \lambda)e_t, \tag{2}$$

denote a general model for this process, where Ω_t is a conditioning information set, $\sigma(\Omega, \lambda)$ the conditional standard deviation of Y_t, and e_t is a zero mean unit variance error process independent of the elements of Ω_t. If $\mu_e = 0$, then tests for conditional asymmetry involve the following null:

$$H_0 : f(e) = f(-e) \text{ for all } e. \tag{3}$$

Bai & Ng (2001) construct tests based on the empirical distribution of e_t and that of $-e_t$. Belaire-Franch & Peiro (2003) apply this and other tests to the Nelson & Plosser (1982) data updated to include 1988.

3 An Entropy-Based Test of Asymmetry

Granger & Maasoumi (1993) considered a normalization of the Bhattacharya-Matusita-Hellinger measure of dependence given by

$$S_\rho = \frac{1}{2}\int_{-\infty}^{\infty} \left(f_1^{1/2} - f_2^{1/2}\right)^2 \, dx \qquad (4)$$

where $f_1 = f(y)$ is the marginal density of the random variable Y and $f_2 = f(\tilde{y})$ that of $\tilde{Y}$, $\tilde{Y}$ being a rotation of Y about its mean.

We consider a kernel-based implementation of equation (4) for the purposes of testing the null of symmetry. Rather than adopt asymptotic-based testing procedures, we elect to use a resampling approach. We do so mainly because critical values obtained from the asymptotic null distribution do not depend on the bandwidth, while the value of the test statistic depends directly on the bandwidth due in part to the fact that the bandwidth is a quantity which vanishes asymptotically. This is a serious drawback in practice, since the outcome of such asymptotic-based tests tends to be quite sensitive to the choice of bandwidth. This has been noted by a number of authors including Robinson (1991) who noted that "substantial variability in the [test statistic] across bandwidths was recorded", which would be most troubling in applied situations due, in part, to numerous competing approaches for data-driven bandwidth choice (see Jones, Marron & Sheather (1996) for an excellent survey article on bandwidth selection for kernel density estimates).

Consider the sample of size $2T$ given by $z = \{Y_1, \dots, Y_T, \tilde{Y}_1, \dots, \tilde{Y}_T\}$. Though the distribution of Y_t may be asymmetric, the distribution of z is symmetric by construction. We may therefore construct the empirical distribution of equation (4) under the null of symmetry by noting that samples drawn with replacement from z will be symmetric almost surely and recomputing equation (4) for B resamples drawn with replacement from z. Given the set of B statistics computed under the null, we may then compute percentiles and use these as the basis for a test of asymmetry. Alternatively, we can compute empirical power via the proportion of the resampled statistics exceeding the actual statistic.

4 Finite-Sample Behavior

We now consider the finite-sample performance of the kernel-based implementation of the test. For what follows, we set the number of bootstrap replications to $B = 99$, and let the sample size assume values $n = 50, 100, 200$. The bandwidth is selected via likelihood cross-validation (Silverman (1986, page 52)) which produces density estimators which are "optimal" according to the Kullback-Leibler criterion. Should one wish to use one of the many alternative methods of bandwidth selection (e.e. see Jones et al. (1996)), one may do so at this stage with no loss of generality.

We consider four DGP's, $N(120, 240)$, $\chi^2(120)$, $\chi^2(80)$, $\chi^2(40)$, $\chi^2(20)$, $\chi^2(10)$, $\chi^2(5)$, and $\chi^2(1)$. Figure 1 plots each DGP to allow the reader to get a sense of the range of distributions considered, from the symmetric $N(\mu, \sigma^2)$ to the range of χ^2 distributions considered.

Table 1 summarizes the finite-sample performance of the proposed test conducted at nominal levels of $\alpha = 0.10, 0.05, 0.01$.

We observe from Table 1 that the test has correct level. Column 2 presents results for the symmetric $N(\mu, \sigma^2)$ distribution for levels $\alpha = 0.10, 0.05, 0.01$. Empirical level does not differ from nominal for any of the sample sizes considered. As the degree of asymmetry increases (moving from column3 through 5), we observe that power increases as it also does when the sample size increases.

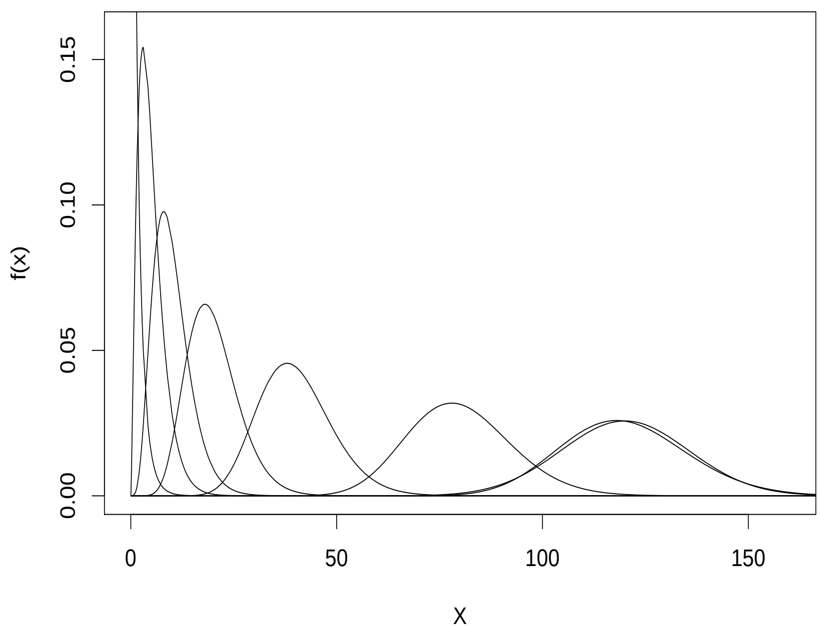

Figure 1: Simulated distributions. The distributions are, from right to left, $N(120, 240)$, $\chi^2(120)$, $\chi^2(80)$, $\chi^2(40)$, $\chi^2(20)$, $\chi^2(10)$, $\chi^2(5)$, and $\chi^2(1)$.

5 Testing for Asymmetry in U.S. Macroeconomic Time Series

Tables 2 and 3 present results for the proposed S_ρ test for unconditional and conditional asymmetry for the updated Nelson & Plosser (1982) data. For the unconditional series we simply rotate the series about its mean, while for the conditional series we employ an AR(P) process with lag order selected via SIC as in Belaire-Franch & Peiro (2003), where $e_t = Y_t - \hat{\delta}_0 - \sum_{j=1}^{p} \hat{\delta}_i Y_{t-j}$. Should the lag order $P = 0$, the test statistics and percentiles for conditional symmetry will be equivalent to those for the unconditional series (ignoring bootstrap resampling error).

Table 1: Empirical rejection frequencies at levels $\alpha = 0.10, 0.05, 0.01$. The degree of asymmetry increases as we go from column 2 (symmetric) to columns 3–5. Column 2 reflects the empirical level of the test, columns 3–5 empirical power.

n	$N(\mu, \sigma^2)$	$\chi^2(120)$	$\chi^2(80)$	$\chi^2(40)$	$\chi^2(20)$	$\chi^2(10)$	$\chi^2(5)$	$\chi^2(1)$
				$\alpha = 0.10$				
50	0.08	0.12	0.21	0.32	0.43	0.58	0.86	0.94
100	0.10	0.22	0.30	0.46	0.68	0.88	0.99	1.00
200	0.09	0.31	0.50	0.73	0.93	1.00	1.00	1.00
				$\alpha = 0.05$				
50	0.04	0.07	0.12	0.18	0.28	0.42	0.70	0.86
100	0.06	0.13	0.22	0.34	0.54	0.81	0.98	1.00
200	0.03	0.19	0.36	0.59	0.86	1.00	1.00	1.00
				$\alpha = 0.01$				
50	0.01	0.02	0.03	0.06	0.12	0.18	0.40	0.66
100	0.03	0.04	0.07	0.17	0.30	0.56	0.85	0.95
200	0.01	0.12	0.18	0.42	0.71	0.96	1.00	1.00

6 Conclusion

We present a simple robust entropy-based test for asymmetry along with a resampling method for obtaining its null distribution. Finite sample performance is examined, while an application to the updated Nelson & Plosser (1982) data indicates power gains relative to the recently proposed test of Bai & Ng (2001).

References

Bai, J. & Ng, S. (2001), ‘A consistent test for conditional symmetry in time series models’, *Journal of Econometrics* **103**, 225–258.

Belaire-Franch, J. & Peiro, A. (2003), ‘Conditional and unconditional asymmetry in u.s. macroeconomic time series’, *Studies in Nonlinear Dynamics and Econometrics* **7**, Article 4.

Table 2: Unconditional symmetry tests and percentiles under the null of symmetry.

Series	$\hat{S}_\rho$	p_{90}	p_{95}	p_{99}
Real GNP	0.030	0.079	0.105	0.168
Nominal GNP	0.108*	0.065	0.079	0.107
Real p/c GNP	0.062	0.088	0.110	0.166
Industrial Production	0.035	0.060	0.075	0.112
Employment Rate	0.024	0.068	0.081	0.116
Unemployment Rate	0.018	0.019	0.024	0.035
GNP Price Deflator	0.165*	0.085	0.098	0.124
CPI	0.234*	0.099	0.116	0.156
Nominal Wages	0.107	0.095	0.115	0.147
Real Wages	0.084	0.134	0.165	0.218
Money Stock	0.043	0.073	0.089	0.126
Velocity	0.162*	0.024	0.030	0.041
Bond Yields	0.345*	0.099	0.117	0.159
S&P 500	0.280*	0.051	0.059	0.081

Crum, W. L. (1923), 'Cycles of rates on commercial paper', *Review of Economics and Statistics* **5**, 17–29.

Granger, C. & Maasoumi, E. (1993), 'A metric entropy measure', *Unpublished notes, Department of Economics, UCSD and SMU* .

Jones, M. C., Marron, J. S. & Sheather, S. J. (1996), 'A brief survey of bandwidth selection for density estimation', *Journal of the American Statistical Association* **91**.

Keynes, J. M. (1936), *The General Theory of Employment, Interest, and Money*, Macmillan Cambridge University Press.

McQueen, G. & Thorley, S. (1993), 'Asymmetric business cycle turning points', *Journal of Monetary Economics* **31**, 341–362.

Mitchell, W. C. (1927), *Business Cycles: The Problem and its Setting*, National Bureau of Economic Research.

Table 3: Conditional symmetry tests and Percentiles under the null of symmetry.

Series	AR(P)	$\hat{S}_\rho$	p_{90}	p_{95}	p_{99}
Real GNP	1	0.015	0.017	0.021	0.032
Nominal GNP	1	0.043	0.055	0.071	0.112
Real p/c GNP	1	0.014	0.017	0.024	0.037
Industrial Production	0	0.035	0.071	0.086	0.115
Employment Rate	3	0.018	0.023	0.030	0.040
Unemployment Rate	3	0.001	0.012	0.017	0.028
GNP Price Deflator	1	0.005	0.027	0.034	0.054
CPI	5	0.014	0.024	0.030	0.042
Nominal Wages	1	0.002	0.026	0.034	0.049
Real Wages	0	0.084	0.140	0.174	0.242
Money Stock	1	0.006	0.017	0.022	0.033
Velocity	0	0.162*	0.021	0.027	0.038
Bond Yields	0	0.345*	0.097	0.117	0.153
S&P 500	0	0.280*	0.051	0.060	0.075

Nelson, C. R. & Plosser, C. (1982), 'Trends and random walks in macroeconomic time series: some evidence and implications', *Journal of Monetary Economics* **10**, 139162.

Robinson, P. M. (1991), 'Consistent nonparametric entropy-based testing', *Review of Economic Studies* **58**, 437–453.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall.

Timmermann, A. & Perez-Quiros, G. (2001), 'Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities', *Journal of Econometrics* **103**, 259–306.

A Series and Residual Plots

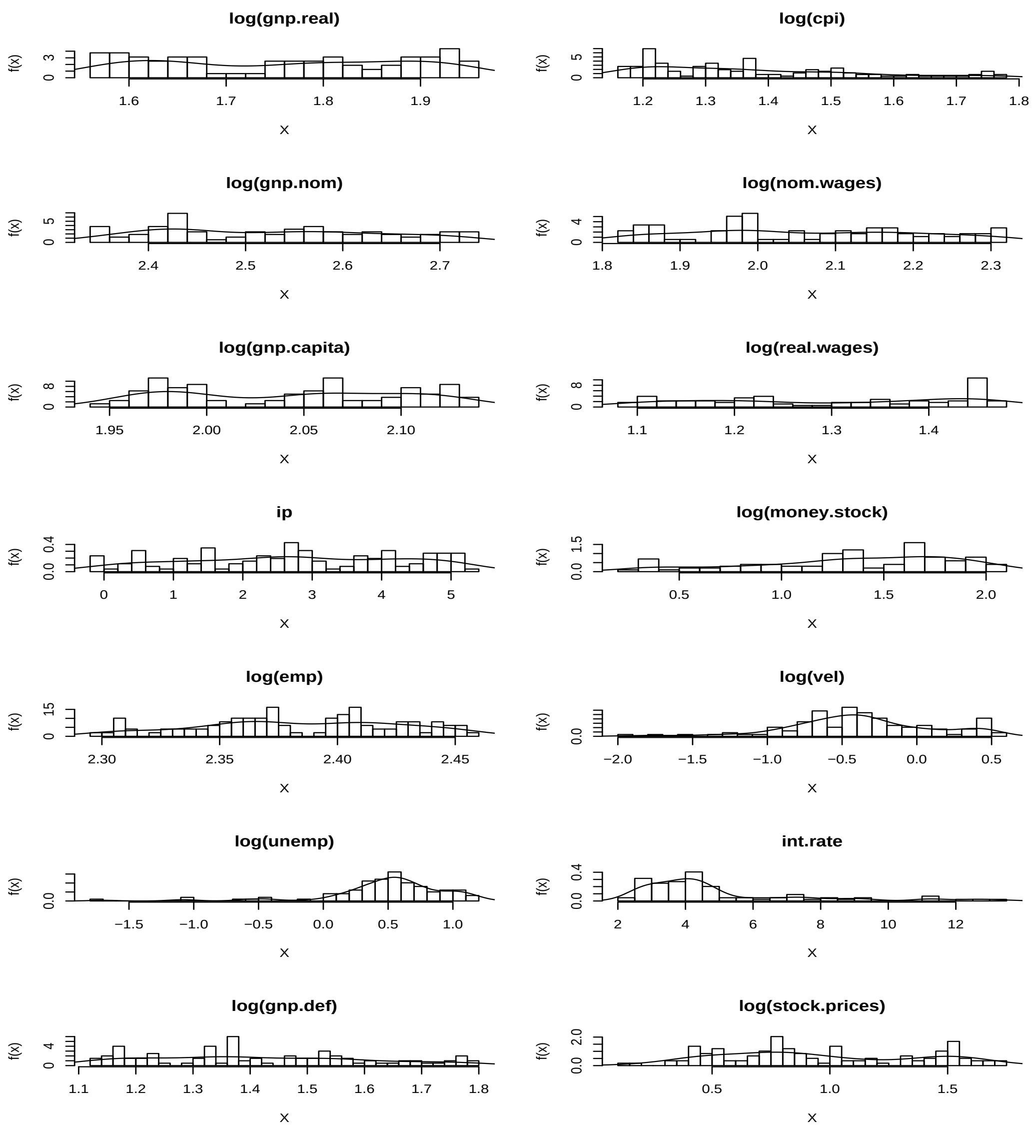

Figure 2: Raw series

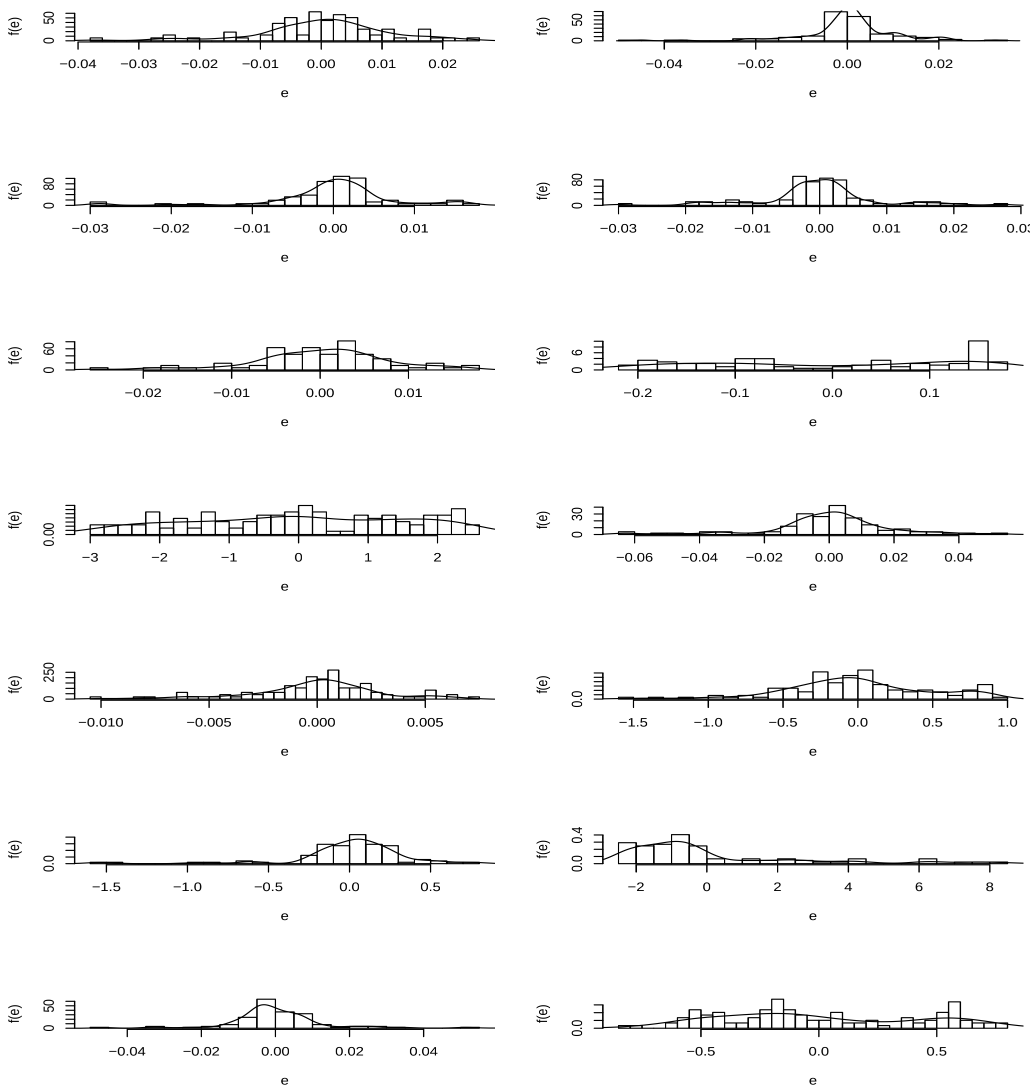

Figure 3: Series residuals: SIC AR(P), no differencing

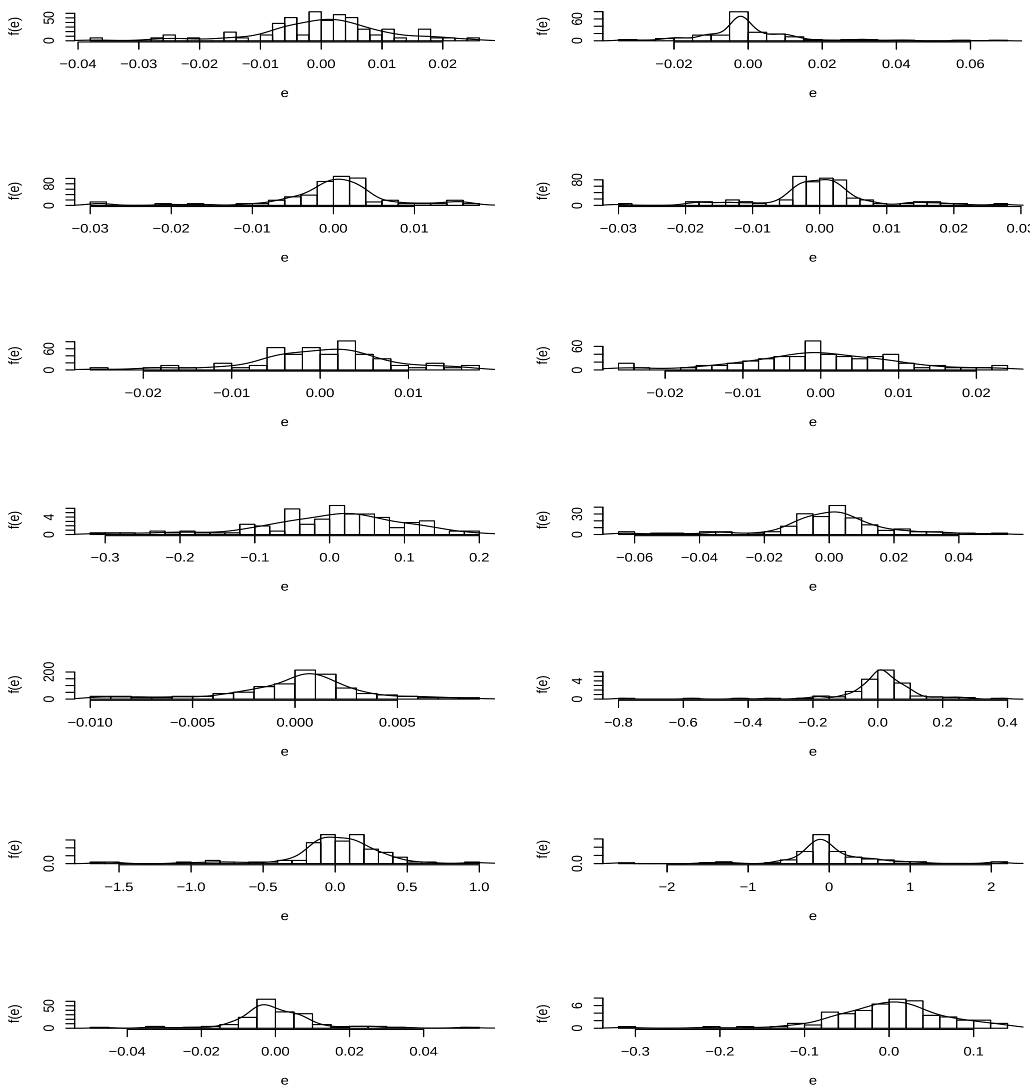

Figure 4: Series residuals: AR(1), no differencing

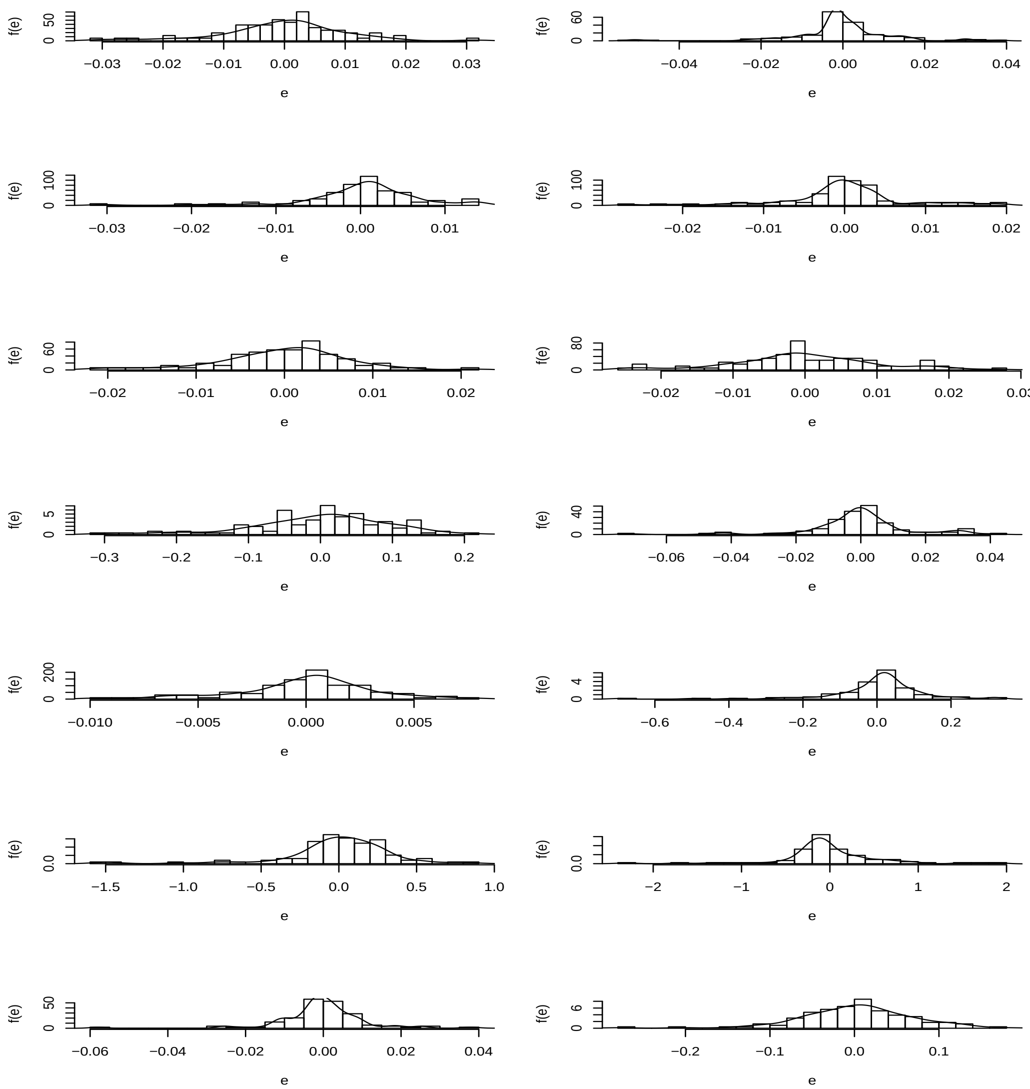

Figure 5: Series residuals: AR(2), no differencing