The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$115	1-year subscription	$145
2-year subscription	$210	2-year subscription	$270
3-year subscription	$285	3-year subscription	$375
1-year student subscription	$ 85	1-year student subscription	$115
1-year institutional subscription	$345	1-year institutional subscription	$375
2-year institutional subscription	$625	2-year institutional subscription	$685
3-year institutional subscription	$875	3-year institutional subscription	$965
Electronic only		**Electronic only**	
1-year subscription	$ 85	1-year subscription	$ 85
2-year subscription	$155	2-year subscription	$155
3-year subscription	$215	3-year subscription	$215
1-year student subscription	$ 55	1-year student subscription	$ 55

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The Stata Journal (2015)
15, Number 3, pp. 822–832

Tests for normality in linear panel-data models

Javier Alejo
Center for Distributive, Labor and Social Sciences
Facultad de Ciencias Económicas
Universidad de La Plata
The National Scientific and Technical
Research Council (CONICET)
La Plata, Argentina
javier.alejo@depeco.econo.unlp.edu.ar

Antonio Galvao
University of Iowa
Iowa City, IA
antonio-galvao@uiowa.edu

Gabriel Montes-Rojas
Universidad de San Andrés-CONICET
Victoria, Argentina
gmontesrojas@udesa.edu.ar

Walter Sosa-Escudero
Universidad de San Andrés-CONICET
Victoria, Argentina
wsosa@udesa.edu.ar

Abstract. We propose a new command, `xtsktest`, for explaining nonnormalities in linear panel-data models. The command performs tests to explore skewness and excess kurtosis, allowing researchers to identify departures from Gaussianity in both error components of a standard panel regression, separately or jointly. The tests are based on recent results by Galvao et al. (2013, *Journal of Multivariate Analysis* 122: 35–52) and extend the classical Jarque–Bera normality test for the case of panel data.

Keywords: st0406, xtsktest, skewness, kurtosis, normality, panel data

1 Introduction

The need to check for nonnormal errors in regression models obeys both to methodological and conceptual reasons. From a strictly methodological viewpoint, lack of Gaussianity sometimes harms the reliability of simple estimation and testing procedures and calls for either better methods under alternative distributional assumptions or robust alternatives whose advantages do not depend on distributional features. Additionally, whether errors should be more appropriately captured by skewed or leptokurtic distributions may be a statistically relevant question.

The normality assumption also plays a crucial role in the validity of inference procedures, specification tests, and forecasting. In the panel-data literature, Blanchard and Mátyás (1996) examine the consequences of nonnormal error components for the performance of several tests. Montes-Rojas and Sosa-Escudero (2011) show that nonnormalities severely affect the performance of the panel-heteroskedasticity tests by Holly and Gardiol (2000) and Baltagi, Bresson, and Pirotte (2006). Despite these concerns, the Gaussian framework is widely used for specification tests in the one-way error-components model; see, for instance, the tests for spatial models in panel data by Baltagi, Song, and Koh (2003) and Baltagi et al. (2007).

Although there is much literature on testing for skewness and kurtosis in cross-sectional and time-series data, including Tolga Ergün and Jun (2010), Bai and Ng (2005), and Bera and Premaratne (2001), results for panel-data models are scarce. Unlike in their cross-section or time-series counterparts, in simple error-components models, lack of Gaussianity may arise in more than one component. Thus an additional problem to that of detecting departures from normality is in identifying which component is causing it. Previous work on the subject includes Gilbert (2002), who exploits cross-moments, and Meintanis (2011), who proposes an omnibus-type test for normality in both components jointly, based on empirical characteristic functions.

The new command `xtsktest` implements a battery of tests to identify nonnormalities in standard error-components panel models, and it is based on recent results by Galvao et al. (2013). For standard regression models, the classical Jarque–Bera test (implemented in Stata with `sktest`) is a simple procedure that detects departures from Gaussianity in the form of skewness and excess kurtosis in the regression error term. A natural concern of panel-data models is identifying which error component (if not both) is the source of nonnormalities. The proposed test allows researchers to explore skewness and excess kurtosis in each component separately or jointly. In this context, the proposed procedure can be seen as extending the famous Jarque–Bera tests for simple panel-data models.

In section 2, we review the results of Galvao et al. (2013) and present the tests. In section 3, we describe the `xtsktest` syntax. In section 4, we then illustrate the procedure by applying the new tests to an investment model studied by Fazzari, Hubbard, and Petersen (1988). In section 5, we conclude with practical suggestions on the proper use of the tests.

2 Skewness and kurtosis in the one-way error-components model

Consider the standard panel-data one-way error-components model

$$y_{it} = x_{it}b + u_i + e_{it}, \quad i = 1, \dots, N,\ t = 1, \dots, T \tag{1}$$

where b is a p-vector of parameters and u_i, e_{it}, and x_{it} are copies of random variables u, e, and x, respectively (b does not contain a constant). As usual, the subscript i refers to individual, and t refers to time. Here u_i and e_{it} refer to the individual-specific and to the remainder error component, respectively, both of which have mean zero.

The quantities of interest are each component's skewness,

$$s_u = \frac{E(u^3)}{\{E(u^2)\}^{3/2}} \quad \text{and} \quad s_e = \frac{E(e^3)}{\{E(e^2)\}^{3/2}}$$

and kurtosis,

$$k_u = \frac{E(u^4)}{\{E(u^2)\}^{2}} \quad \text{and} \quad k_e = \frac{E(e^4)}{\{E(e^2)\}^{2}}$$

Galvao et al. (2013) construct statistics for testing for skewness and kurtosis in the individual-specific and the remainder components, separately and jointly. When the underlying distribution is normal, the null hypotheses of interest become $H_0^{s_u}: s_u = 0$ and $H_0^{s_e}: s_e = 0$ for skewness and $H_0^{k_u}: k_u = 3$ and $H_0^{k_e}: k_e = 3$ for kurtosis. Moreover, under normality, the null hypotheses for these cases are given by

$$H_0^{s_u \& k_u}: s_u = 0 \text{ and } k_u = 3$$

$$H_0^{s_e \& k_e}: s_e = 0 \text{ and } k_e = 3$$

The statistics for symmetry are

$$\widehat{SK}_u^{(1)} = \widehat{E(u^3)} = \frac{T^2 - 3T}{T^2 - 3T + 2}\mathbb{E}\left(\overline{\widehat{\epsilon}}_i^{3}\right) - \frac{1}{T^2 - 3T + 2}\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{3}} - 3\overline{\widehat{\epsilon}}_i\overline{\widehat{\epsilon}_i^{2}}\right)$$

and

$$\widehat{SK}_e^{(1)} = \widehat{E(e^3)} = \frac{1}{1 - 3T^{-1} + 2T^{-2}}\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{3}} - 3\overline{\widehat{\epsilon}}_i\overline{\widehat{\epsilon}_i^{2}} + 2\overline{\widehat{\epsilon}}_i^{3}\right)$$

where $\widehat{\epsilon}_{it}$ denotes the ordinary least-squares (OLS) residuals of model (1), a line over a variable with a subscript i indicates a group average with $\overline{\widehat{\epsilon}_i^{j}} = 1/T\sum_{i=1}^{T}\widehat{\epsilon}_{it}^{j}$, $j = 1$, 2, 3, 4, and $\mathbb{E}(W_i) = 1/N\sum_{i=1}^{N}W_i$ for a generic variable W_i indexed by i. These correspond to the statistics $\widehat{\mu}_3$ and $\widehat{\nu}_3$, respectively, in Galvao et al. (2013, 37). The statistics for kurtosis are

$$\begin{aligned}\widehat{KU}_u^{(1)} = \widehat{E(u^4)} - 3\left\{\widehat{E\left(u^2\right)}\right\}^2 &= \mathbb{E}\left(\overline{\widehat{\epsilon}}_i^{4}\right)\frac{T^3 - 4T^2 + 6T}{T^3 - 4T^2 + 6T - 3} \\ &- \frac{\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{4}}\right) - 4\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{3}}\overline{\widehat{\epsilon}}_i\right) + 6\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{2}}\overline{\widehat{\epsilon}}_i^{2}\right)}{T^3 - 4T^2 + 6T - 3} \\ &- \frac{(T-1)(3T^3 - 12T^2 + 12T + 3)}{(T^3 - 4T^2 + 6T - 3)T^3}\widehat{\sigma}_e^4 - \frac{6}{T}\widehat{\sigma}_u^2\widehat{\sigma}_e^2 - 3\widehat{\sigma}_u^4\end{aligned}$$

and

$$\begin{aligned}\widehat{KU}_e^{(1)} &= \frac{\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{4}}\right) - 4\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{3}}\overline{\widehat{u}}_i\right) + 6\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{2}}\overline{\widehat{\epsilon}}_i^{2}\right) - 3\mathbb{E}\left(\overline{\widehat{\epsilon}}_i^{4}\right)}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}} \\ &- \frac{(T-1)(6T^{-2} - 12T^{-3})}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}}\widehat{\sigma}_e^4 - 3\widehat{\sigma}_e^4\end{aligned}$$

where

$$\widehat{\sigma}_e^2 = \frac{1}{1 - T^{-1}}\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{2}}\right) - \frac{1}{1 - T^{-1}}\mathbb{E}\left(\overline{\widehat{\epsilon}}_i^{2}\right)$$

$$\widehat{\sigma}_u^2 = \frac{T}{T-1}\mathbb{E}\left(\overline{\widehat{\epsilon}}_i^{2}\right) - \frac{1}{T-1}\mathbb{E}\left(\overline{\widehat{\epsilon}_i^{2}}\right)$$

These correspond to the zero mean transformation of the statistics $\widehat{\mu}_4$ and $\widehat{\nu}_4$, respectively, in Galvao et al. (2013, 38).

Alternative statistics can also be presented in a standardized way as $\widehat{SK}_u^{(2)} = \widehat{E(u^3)}/[\{\widehat{E(u^2)}\}^{3/2}]$ and $\widehat{SK}_e^{(2)} = \widehat{E(e^3)}/[\{\widehat{E(e^2)}\}^{3/2}]$ for symmetry [(3) and (2), respectively, in Galvao et al. (2013, 37)] and $\widehat{KU}_u^{(2)} = \widehat{E(u^4)}/[\{\widehat{E(u^2)}\}^2] - 3$ and $\widehat{KU}_e^{(2)} = \widehat{E(e^4)}/[\{\widehat{E(e^2)}\}^2] - 3$ [zero mean transformation of the statistics in (5) and (4), respectively, in Galvao et al. (2013, 38)] for kurtosis. Each statistic is consistent and, when properly standardized, follows an $N(0,1)$ asymptotic law under the corresponding null hypothesis. However, each may differ in small samples. Tests for joint symmetry and kurtosis are constructed using $(\widehat{SK}_u^{(j)})^2 + (\widehat{KU}_u^{(j)})^2$ and $(\widehat{SK}_e^{(j)})^2 + (\widehat{KU}_e^{(j)})^2$, $j = 1, 2$, each following a χ^2_2 asymptotic law under the corresponding null hypothesis.

The variance of the statistics depends on the higher-order single- and cross-moments of u and e (up to the sixth for skewness and eighth for kurtosis). Its computation is thus very cumbersome. Moreover, the analytical variance depends on the statistic used for skewness and kurtosis (for example, $\widehat{SK}_e^{(1)}$ or $\widehat{SK}_e^{(2)}$). Direct estimation of the asymptotic variances is possible using an outer product of the gradient strategy, but extensive Monte Carlo experimentation shows that the bootstrap performs better. Following Galvao et al. (2013), we implement the tests using the bootstrap, randomly drawing individuals with replacement while maintaining the unaltered time-series structure to estimate the variances of the skewness and kurtosis test statistics. The `bootstrap` command in Stata offers a flexible and efficient computational framework to implement these tests by specifying the `cluster()` option at the individual level.

Simulation experiments in Galvao et al. (2013) show that the tests are consistent (as $N \to \infty$) and responsive to both deviations in skewness and kurtosis and that deviations in one component do not affect the empirical size in the other component, thus allowing one to identify the source of skewness and kurtosis in each error component.

3 The xtsktest command

3.1 Syntax

`xtsktest` [*varlist*] [*if*] [, `reps(`#`)` `seed(`#`)` `standard`]

3.2 Options

`reps(`#`)` specifies the number of bootstrap replications. The default is `reps(50)`.

`seed(`#`)` specifies the seed for the random-number generator in the bootstrap procedure; see [R] **set seed**.

`standard` specifies whether the skewness and kurtosis statistics are standardized by the estimated variance. The default is no standardization.

3.3 Remarks

`xtsktest` can be used both as a standard command or as a postestimation command after an OLS or random-effects model (see [R] **regress** and [XT] **xtreg**). In the former, the command requires at least one variable in the *varlist*, while in the latter, *varlist* is not required. Example 1 shows the former; examples 2 and 3 use `xtsktest` as a postestimation command.

3.4 Stored results

`xtsktest` stores the following in `e()`:

Matrices

`e(xtsk_test)`	skewness and kurtosis test results, one per row; first column for point estimation, second for standard errors, and third for p-values
`e(joint_test)`	joint skewness and kurtosis test results, one per row; first column for chi-squared statistics and second for p-values

4 Empirical application: Investment equation

In this section, we apply the developed tests to the Fazzari, Hubbard, and Petersen (1988) investment equation model, where a firm's investment is regressed on an observed measure of investment demand (Tobin's q) and cash flow. This is one of the most well-known models in the corporate investment literature, and we use this application to illustrate our theoretical results. As a result of Fazzari, Hubbard, and Petersen (1988), investment–cash-flow sensitivities became a standard metric for examining the impact of financing imperfections on corporate investment (Stein 2003).

Tobin's q is the ratio of the market valuation of a firm and the replacement value of its assets. A high value of q for a firm indicates an attractive investment opportunity, whereas a low value of q indicates the opposite. Investment theory is also interested in the effect of cash flow, because the theory predicts that financially constrained firms are more likely to rely on internal funds to finance investment (see, for example, Erickson and Whited [2000]). The baseline model in the literature is

$$I_{it}/K_{it} = \alpha + \beta q_{it-1} + \gamma CF_{it-1}/K_{it-1} + u_i + e_{it}$$

where I denotes investment, K denotes capital stock, CF denotes cash flow, q denotes Tobin's q, u represents the firm-specific effect, and e is the innovation term.

We check for skewness and kurtosis in both u and e using the proposed tests. We are interested in testing for skewness and kurtosis for at least three reasons. First, testing normality plays a key role in forecasting models at the firm level. Second, asymmetry in both components is used for solving measurement-error problems in Tobin's q. The operationalization of q is not clear-cut, so estimation poses a measurement-error problem. Many empirical investment studies found the q theory of investment to perform poorly, although this theory has a good performance when measurement error is purged as in Erickson and Whited (2000). Their method requires asymmetry in the error term to identify the effect of q on firm investment. Third, skewness and kurtosis by themselves provide information about the industry investment patterns. Skewness in u determines that a few firms either invest or disinvest considerably more than the rest, while kurtosis in u determines that a few firms locate at both sides of the investment line—that is, some invest a large amount, while others disinvest a large amount. Skewness or kurtosis in e shows that the large values of investment correspond to firm-level shocks.

We follow Almeida, Campello, and Galvao (2010), who considered a sample of manufacturing firms (standard industrial classifications 2,000 to 3,999) from 2000–2005 with data from Compustat's PST and Full Coverage files. Only firms with observations in every year are used to construct a balanced panel of firms for the five-year period. Moreover, following those authors, we eliminate firms for which cash holdings exceeded the value of total assets and those displaying asset or sales growth exceeding 100%. Our final sample consists of 410 firm-years and 82 firms. Because we consider only the firms that report information in each of the five years, the sample consists mainly of relatively large firms.

To demonstrate the use of `xtsktest` in this case, we must first open the dataset and declare it to be panel data; see [XT] **xtset**.

```
. version 13
. use investment.dta
. xtset idcode time
       panel variable:  idcode (strongly balanced)
        time variable:  time, 2 to 6
                delta:  1 unit
```

First, we consider an OLS estimation of the effect of Tobin's q and cash flows on investment.

```
. regress investment tobinq cashflow

      Source |       SS           df       MS      Number of obs   =       410
-------------+----------------------------------   F(2, 407)       =     89.07
       Model |  .536747282         2  .268373641   Prob > F        =    0.0000
    Residual |  1.22632448       407  .003013082   R-squared       =    0.3044
-------------+----------------------------------   Adj R-squared   =    0.3010
       Total |  1.76307176       409  .004310689   Root MSE        =    .05489

------------------------------------------------------------------------------
  investment |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      tobinq |   .0384663   .0094022     4.09   0.000     .0199834    .0569492
    cashflow |   .1117721   .0096142    11.63   0.000     .0928724    .1306718
       _cons |   .0669764   .0087876     7.62   0.000     .0497016    .0842512
------------------------------------------------------------------------------
```

Second, we consider a one-way error-components random-effects model.

```
. xtreg investment tobinq cashflow, re

Random-effects GLS regression                   Number of obs      =       410
Group variable: idcode                          Number of groups   =        82

R-sq:                                           Obs per group:
     within  = 0.1014                                         min =          5
     between = 0.3583                                         avg =        5.0
     overall = 0.2779                                         max =          5

                                                Wald chi2(2)       =     84.09
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
  investment |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      tobinq |   .0673706   .0129138     5.22   0.000       .04206    .0926812
    cashflow |   .0824715   .0115191     7.16   0.000     .0598944    .1050486
       _cons |   .0516002   .0127921     4.03   0.000     .0265281    .0766722
-------------+----------------------------------------------------------------
     sigma_u |   .0380806
     sigma_e |  .03857635
         rho |  .49353308   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

The results show a positive and significant effect of both Tobin's q and cash flows on investment flows in both models. The random-effects model also shows that there is considerable variation across firms in terms of unobservables. Half the variation is due to the firm-specific component u_i, and the other half is due to the remainder component e_{it}. Note that the presence of firm-specific effects determines that OLS standard errors are not correct, while the random effects are.

Here we consider the use of `xtsktest` as an estimation command of the skewness and kurtosis of each component. We can implement the command in the following three equivalent ways: as a single command (example 1), as a postestimation command after OLS, or as a postestimation command after random effects. We consider the implementation with 500 bootstrap replications and with a random-number seed (= 123) (default options have 50 bootstrap replications and no random-number seed).

```
. * Example 1, command mode
. xtsktest investment tobinq cashflow, reps(500) seed(123)
(running _xtsktest_calculations on estimation sample)

Bootstrap replications (500)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
..................................................   450
..................................................   500

Tests for skewness and kurtosis                  Number of obs     =        410
                                                 Replications      =        500

                                      (Replications based on 82 clusters in idcode)

------------------------------------------------------------------------------
             |  Observed   Bootstrap                         Normal-based
             |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  Skewness_e |  .0000387   .0000137     2.81   0.005      .0000117    .0000656
  Kurtosis_e |  9.33e-06   1.92e-06     4.87   0.000      5.58e-06    .0000131
  Skewness_u |  .0000511   .0000171     2.99   0.003      .0000176    .0000847
  Kurtosis_u |  4.27e-08   1.27e-06     0.03   0.973     -2.44e-06    2.53e-06
------------------------------------------------------------------------------

Joint test for Normality on e:        chi2(2) =  31.64     Prob > chi2 = 0.0000
Joint test for Normality on u:        chi2(2) =   8.93     Prob > chi2 = 0.0115
------------------------------------------------------------------------------
```

The output shows the observed coefficients of the four statistics (without standardization, $\widehat{SK}_e^{(1)} = 0.0000387$, $\widehat{KU}_e^{(1)} = 9.33e-06$, $\widehat{SK}_u^{(1)} = 0.0000511$, and $\widehat{KU}_u^{(1)} = 4.27e-08$) used for symmetry and kurtosis for each error component in the first column. The next columns show the standard errors computed by bootstrap replications, the z statistics, the p-values, and the 95% confidence intervals using the normal approximation. Finally, the lower part of the output shows the joint test for normality on each component of the error term and the respective p-values. The tests reveal that both components are asymmetric (with right symmetry), while only the remainder component e has excess kurtosis. Thus, while we expect the occurrence of large positive investment shocks [$E(e^3) > 0$], these are systematic in some firms [that is, $E(u^3) > 0$]. Asymmetry thus produces the rejection of the null hypothesis of normality in both error components, although the rejection is stronger for the remainder than for the firm-specific component.

We also evaluate symmetry and kurtosis in each component using the standardized statistics, $\widehat{SK}_e^{(2)}$, $\widehat{KU}_e^{(2)}$, $\widehat{SK}_u^{(2)}$, and $\widehat{KU}_u^{(2)}$. These can be implemented with the option `standard`.

```
. * Example 2, standardized coefficients
. xtsktest investment tobinq cashflow, reps(500) seed(123) standard
(running _xtsktest_calculations on estimation sample)
Bootstrap replications (500)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
..................................................   450
..................................................   500
Tests for skewness and kurtosis                 Number of obs     =        410
                                                Replications      =        500
                                    (Replications based on 82 clusters in idcode)
```

	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
Skewness_e	.6040947	.1523494	3.97	0.000	.3054954	.902694
Kurtosis_e	3.645848	.6932803	5.26	0.000	2.287044	5.004653
Skewness_u	.9857612	.2176725	4.53	0.000	.5591309	1.412391
Kurtosis_u	.0220666	.4963144	0.04	0.965	-.9506917	.994825

```
Joint test for Normality on e:        chi2(2) =  43.38    Prob > chi2 = 0.0000
Joint test for Normality on u:        chi2(2) =  20.51    Prob > chi2 = 0.0000

Note: standardized coefficients
```

As expected, the results do not differ from those presented with the nonstandardized statistics. The numeric results, however, provide an easier interpretation of the excess kurtosis in the remainder component with a value of $\widehat{KU}_e^{(2)} = 3.645848$ and the firm-specific component $\widehat{KU}_u^{(2)} = 0.0220666$. The joint test for normality in u, however, provides a higher chi-squared value with a clearer rejection than in the previous examples using nonstandardized coefficients.

5 Conclusion

In this article, we implemented tests for skewness and symmetry and kurtosis of the error components in linear panel-data random-effects models. `xtsktest` allows one to evaluate each error component's third and fourth moments. This can be used as an alternative to the Jarque–Bera test in panel-data models.

As previously discussed, checking for skewness and kurtosis in the error components plays an important role in testing and estimation in linear panel-data models. Deviations from symmetry and kurtosis of three invalidate methods that are not robust to normality. Moreover, estimating third and fourth moments is also important for forecasting in panel-data models (see Baltagi [2008] for a discussion).

6 References

Almeida, H., M. Campello, and A. F. Galvao, Jr. 2010. Measurement errors in investment equations. *Review of Financial Studies* 23: 3279–3328.

Bai, J., and S. Ng. 2005. Tests for skewness, kurtosis, and normality for time series data. *Journal of Business and Economic Statistics* 23: 49–60.

Baltagi, B. H. 2008. Forecasting with panel data. *Journal of Forecasting* 27: 153–173.

Baltagi, B. H., G. Bresson, and A. Pirotte. 2006. Joint LM test for homoskedasticity in a one-way error component model. *Journal of Econometrics* 134: 401–417.

Baltagi, B. H., S. H. Song, B. C. Jung, and W. Koh. 2007. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics* 140: 5–51.

Baltagi, B. H., S. H. Song, and W. Koh. 2003. Testing panel data regression models with spatial error correlation. *Journal of Econometrics* 117: 123–150.

Bera, A. K., and G. Premaratne. 2001. Adjusting the tests for skewness and kurtosis for distributional misspecifications. Working paper, University of Illinois at Urbana-Champaign. http://www.business.uiuc.edu/Working_Papers/papers/01-0116.pdf.

Blanchard, P., and L. Mátyás. 1996. Robustness of tests for error components models to non-normality. *Economics Letters* 51: 161–167.

Erickson, T., and T. M. Whited. 2000. Measurement error and the relationship between investment and q. *Journal of Political Economy* 108: 1027–1057.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen. 1988. Financing constraints and corporate investment. *Brookings Papers on Economic Activity* 1: 141–195.

Galvao, A. F., G. Montes-Rojas, W. Sosa-Escudero, and L. Wang. 2013. Tests for skewness and kurtosis in the one-way error component model. *Journal of Multivariate Analysis* 122: 35–52.

Gilbert, S. 2002. Testing the distribution of error components in panel data models. *Economics Letters* 77: 47–53.

Holly, A., and L. Gardiol. 2000. A score test for individual heteroscedasticity in a one-way error components model. In *Panel Data Econometrics: Future Directions*, ed. J. Krishnakumar and E. Ronchetti, 199–211. Amsterdam: Elsevier.

Meintanis, S. G. 2011. Testing for normality with panel data. *Journal of Statistical Computation and Simulation* 81: 1745–1752.

Montes-Rojas, G., and W. Sosa-Escudero. 2011. Robust tests for heteroskedasticity in the one-way error components model. *Journal of Econometrics* 160: 300–310.

Stein, J. C. 2003. Agency, information and corporate investment. In *Handbook of the Economics of Finance: Volume 1A—Corportate Finance*, ed. G. M. Constantinides, M. Harris, and R. M. Stulz, 111–163. Amsterdam: Elsevier.

Tolga Ergün, A., and J. Jun. 2010. Conditional skewness, kurtosis, and density specification testing: Moment-based versus nonparametric tests. *Studies in Nonlinear Dynamics and Econometrics* 14.

About the authors

Javier Alejo is a senior researcher who specializes in labor economics and inequality.

Antonio Galvao is an associate professor who specializes in econometric theory and statistics.

Gabriel Montes-Rojas is a senior researcher who specializes in econometric theory and statistics.

Walter Sosa-Escudero is an associate professor who specializes in econometric theory and statistics.