Tests for skewness and kurtosis in the one-way error components model

Antonio F. Galvao* Gabriel Montes-Rojas† Walter Sosa-Escudero‡
Liang Wang§

July 2, 2013

Abstract

This paper derives tests for skewness and kurtosis for the panel data one-way error components model. The test statistics are based on the between and within transformations of the pooled OLS residuals, and are derived in a moment conditions framework. We establish the limiting distribution of the test statistics for panels with large cross-section and fixed time-series dimension. The tests are implemented in practice using the bootstrap. The proposed methods are able to detect departures away from normality in the form of skewness and kurtosis, and to identify whether these occur at the individual, remainder, or both error components. The finite sample properties of the tests are studied through extensive Monte Carlo simulations, and the results show evidence of good finite sample performance.

Key Words: Panel data, Error components, Skewness, Kurtosis, Normality

AMS subject classifications: 62F03, 62F05, 62H15

*Corresponding author: Department of Economics, University of Iowa, W210 Pappajohn Business Building, 21 E. Market Street, Iowa City, IA 52242. E-mail: antonio-galvao@uiowa.edu

†Department of Economics, City University London, Northampton Square, London EC1V 0HB, UK. E-mail: Gabriel.Montes-Rojas.1@city.ac.uk

‡Department of Economics, Universidad de San Andrés, Argentina and CONICET. E-mail: wsosa@udesa.edu.ar

§Department of Economics, University of Wisconsin-Milwaukee, NWQ B, 2025 E. Newport Ave., Milwaukee, WI 53201. E-mail: wang42@uwm.edu

1 Introduction

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

The normality assumption also plays a crucial role in the validity of specification tests. Blanchard and Mátyás (1996) examine the consequences of non-normal error components for the performance of several tests. In a recent application, Montes-Rojas and Sosa-Escudero (2011) show that non-normalities severely affect the performance of the panel heteroskedasticity tests by Holly and Gardiol (2000) and Baltagi, Bresson and Pirotte (2006), in line with the results of Evans (1992) for the cross-sectional case. 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, Song, Jung, and Koh (2007).

Even though there is a large literature on testing for skewness and kurtosis in cross-sectional and time-series data, including Ergun and Jun (2010), Bai and Ng (2005), Premaratne and Bera (2005), Dufour, Khalaf and Beaulieu (2003), Bera and Premaratne (2001), Henze (1994) and Lutkepohl and Theilen (1991) to cite a few of an extensive list that dates back to the seminal article by Jarque and Bera (1981), results for panel data models are scarce. A natural complication is that, unlike 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 away from normality is the identification of which component is causing it. Previous work on the subject include 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.

This paper develops tests for skewness (lack of symmetry), kurtosis, and normality for panel data one-way error component models. The tests are constructed based on moment conditions of the within and between transformations of the OLS residuals. These conditions are exploited to develop tests for skewness and kurtosis in the individual-specific and the remainder components, separately and jointly. We show that under the corresponding null hypothesis the limiting distributions of the tests are asymptotically normal. To obtain the asymptotic distributions of the test statistics, we consider the most important case where the number of individuals, N, goes to infinity, but the number of time periods, T, is fixed and might be small. The proposed methods and associated limiting theory are important in practice because, in the panel data case, the standard Bera-Jarque test is not able to disentangle the departures of the individual and remainder components from non-Gaussianity.

The proposed tests are implemented in practice using a bootstrap procedure. Since the tests are asymptotically normal, the bootstrap can be used to compute the corresponding variance-covariance matrices of the statistics of interest and carry out inference. In particular, the tests are implemented using a cross-sectional bootstrap. We formally prove the consistency of the bootstrap method applied to our case of short panels.

A Monte Carlo study is conducted to assess the finite sample performance of the tests in terms of size and power. The Monte Carlo simulations show that the proposed tests and their bootstrap implementation work well for both skewness and kurtosis, even in small samples similar to those used in practice. The results confirm that the test for the individual specific component depends on the cross-section dimension only, and hence it is invariant to the time-series dimension. The proposed tests detect departures away from the null hypothesis of skewness and/or kurtosis in each component, and are robust to the presence of skewness and/or kurtosis in the other component.

Finally, to highlight the usefulness of the proposed tests, we apply the new tests to the Fazzari, Hubbard and Petersen (1988) investment equation model, in which firm investment is regressed on a proxy for investment demand (Tobin's q) and cash flow.

The paper is organized as follows. Section 2 presents the relevant moment conditions that characterize skewness and kurtosis for each error component. Section 3 derives the tests and their asymptotic distributions. Section 4 describes the implementation through bootstrap. Section 5 presents Monte Carlo results. A brief application is given in Section 6. Finally, Section 7 concludes and discusses extensions.

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

2.1 The model and the hypotheses

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

$$y_{it} = \alpha_0 + \boldsymbol{x}_{it}^{\top}\boldsymbol{\beta}_0 + u_{it},\ u_{it} = \mu_i + \nu_{it}, \quad i = 1, ...N,\ t = 1, ..., T, \tag{1}$$

where α_0 is a constant, $\boldsymbol{\beta}_0$ is a p-vector of parameters, and μ_i, ν_{it}, and $\boldsymbol{x}_{it}$ are copies of random variables μ, ν, and $\boldsymbol{x}$, respectively. As usual, the subscript i refers to individual and t to time. Here μ_i and ν_{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 skewness,

$$s_\mu = \frac{\mu_3}{\sigma_\mu^3} = \frac{\mathrm{E}[\mu^3]}{(\mathrm{E}[\mu^2])^{3/2}}, \quad \text{and} \quad s_\nu = \frac{\nu_3}{\sigma_\nu^3} = \frac{\mathrm{E}[\nu^3]}{(\mathrm{E}[\nu^2])^{3/2}},$$

and kurtosis,

$$k_\mu = \frac{\mu_4}{\sigma_\mu^4} = \frac{\mathrm{E}[\mu^4]}{(\mathrm{E}[\mu^2])^{2}}, \quad \text{and} \quad k_\nu = \frac{\nu_4}{\sigma_\nu^4} = \frac{\mathrm{E}[\nu^4]}{(\mathrm{E}[\nu^2])^{2}}.$$

We are interested in 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 $H_0^{s_\mu} : s_\mu = 0$ and $H_0^{s_\nu} : s_\nu = 0$ for skewness and $H_0^{k_\mu} : k_\mu = 3$ and $H_0^{k_\nu} : k_\nu = 3$ for kurtosis. We also consider testing for skewness and kurtosis jointly. Under normality, the null hypotheses for these cases are given by

$$H_0^{s_\mu \& k_\mu} : s_\mu = 0 \text{ and } k_\mu = 3,$$

$$H_0^{s_\nu \& k_\nu} : s_\nu = 0 \text{ and } k_\nu = 3.$$

It is common in the statistics and econometrics literature to check for the third and fourth moment of a random variable and compare them with the corresponding values of a normal distribution. This corresponds to a test for normality of each component. Thus, the last hypotheses can be regarded as tests for normality. The following sections develop the corresponding test statistics.

We introduce the following notation. $\overline{u}_i \equiv T^{-1}\sum_{t=1}^{T} u_{it}$ are the *between* residuals, and $\widetilde{u}_{it} \equiv u_{it} - \overline{u}_i$ are the *within* residuals. Let $\overline{\nu}_i \equiv T^{-1}\sum_{t=1}^{T} \nu_{it}$ such that $\overline{u}_i = \mu_i + \overline{\nu}_i$.

Similarly, $\overline{\boldsymbol{x}}_i \equiv \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_{it}$. In general, a line over a variable with a subscript i indicates a *group* average. A tilde will denote variables expressed as deviations from the corresponding group mean. A line over a variable without a subscript indicates the total average across i and t, e.g., $\overline{\boldsymbol{x}} \equiv \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{x}_{it}$. Also, for a generic random vector X_i indexed by i, let $\mathbb{E}[X_i]$ denote $\frac{1}{N}\sum_{i=1}^{N}X_i$. Let $\widehat{u}_{it} = y_{it} - \widehat{\alpha}_0 - \boldsymbol{x}_{it}^{\top}\widehat{\boldsymbol{\beta}}$ be the pooled ordinary least squares (OLS) residuals, where $\widehat{\boldsymbol{\beta}}$ is the pooled OLS estimator. It is straightforward to show that the slope coefficients admit an asymptotic representation such that $\sqrt{N}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) = O_p(1)$.

2.2 Skewness

In this section we develop statistics for testing skewness. In order to derive moment conditions to test for skewness in each component, consider the cube of the *between* residuals, $\overline{u}_i^3 = (\mu_i + \overline{\nu}_i)^3 = \mu_i^3 + \overline{\nu}_i^3 + 3\mu_i^2\overline{\nu}_i + 3\mu_i\overline{\nu}_i^2$. Taking expectations on both sides of the last equation, we have

$$\mathrm{E}[\overline{u}_i^3] = \mu_3 + T^{-2}\nu_3.$$

The cube of the *within* residuals is given by $\widetilde{u}_{it}^3 = (\nu_{it} - \overline{\nu}_i)^3 = \nu_{it}^3 - \overline{\nu}_i^3 - 3\nu_{it}^2\overline{\nu}_i + 3\nu_{it}\overline{\nu}_i^2$, and after taking expectations

$$\mathrm{E}[\widetilde{u}_{it}^3] = \nu_3 - T^{-2}\nu_3 - 3T^{-1}\nu_3 + 3T^{-2}\nu_3 = \nu_3\left(1 - 3T^{-1} + 2T^{-2}\right).$$

Solving these equations for ν_3 and μ_3, we obtain

$$\begin{aligned}
\nu_3 &= \frac{1}{1 - 3T^{-1} + 2T^{-2}}\mathrm{E}[\widetilde{u}_{it}^3],\\
\mu_3 &= \mathrm{E}[\overline{u}_i^3] - \frac{1}{T^2 - 3T + 2}\mathrm{E}[\widetilde{u}_{it}^3].
\end{aligned}$$

Therefore, the tests for skewness in each component make use of $\overline{u}_i^3$ and $\widetilde{u}_{it}^3$. To carry tests in practice, we need to construct the sample counterpart of these equations. First, we rewrite the equations. Recall that $\widetilde{u}_{it}^3 = (u_{it} - \overline{u}_i)^3 = u_{it}^3 - \overline{u}_i^3 - 3u_{it}^2\overline{u}_i + 3u_{it}\overline{u}_i^2$, and that in the one-way error components model (using the assumptions below), for each i, the $\{\widetilde{u}_{it}\}$'s are identically distributed, then it follows that $\mathrm{E}[\widetilde{u}_{it}^3] = \mathrm{E}[\overline{\widetilde{u}_i^3}]$ where $\overline{\widetilde{u}_i^3} = \frac{1}{T}\sum_{t=1}^{T}\widetilde{u}_{it}^3$, and

$$\begin{aligned}
\nu_3 &= \frac{1}{1 - 3T^{-1} + 2T^{-2}}\mathrm{E}[\overline{u_i^3} - 3\overline{u}_i\overline{u_i^2} + 2\overline{u}_i^3],\\
\mu_3 &= \frac{T^2 - 3T}{T^2 - 3T + 2}\mathrm{E}[\overline{u}_i^3] - \frac{1}{T^2 - 3T + 2}\mathrm{E}[\overline{u_i^3} - 3\overline{u}_i\overline{u_i^2}].
\end{aligned}$$

However, the above equations are functions of the unobservable innovations. Thus, we define $\widehat{\nu}_3$ and $\widehat{\mu}_3$ as the estimators of ν_3 and μ_3, respectively, which are constructed using $\widehat{u}_{it}$, the OLS residuals. Then, we have that

$$\begin{aligned}\widehat{\nu}_3 =& \frac{1}{1-3T^{-1}+2T^{-2}}\mathbb{E}[\overline{\widehat{u}_i^3} - 3\overline{\widehat{u}}_i\overline{\widehat{u}_i^2} + 2\overline{\widehat{u}}_i^3],\\ \widehat{\mu}_3 =& \frac{T^2-3T}{T^2-3T+2}\mathbb{E}[\overline{\widehat{u}}_i^3] - \frac{1}{T^2-3T+2}\mathbb{E}[\overline{\widehat{u}_i^3} - 3\overline{\widehat{u}}_i\overline{\widehat{u}_i^2}].\end{aligned}$$

Finally, the following statistics are useful to construct tests for skewness

$$\widehat{SK}_\nu = \frac{\widehat{\nu}_3}{\widehat{\sigma}_\nu^3}, \tag{2}$$

$$\widehat{SK}_\mu = \frac{\widehat{\mu}_3}{\widehat{\sigma}_\mu^3}. \tag{3}$$

In Section 3 we construct tests for skewness based on these statistics, and derive their limiting distributions under some regularity conditions (Assumptions 1 and 2, in the next section). The estimators $\widehat{\sigma}_\nu^3$ and $\widehat{\sigma}_\mu^3$ are given in Appendix A1. It is easy to show that under $H_0^{s_\nu}$, $\widehat{SK}_\nu \xrightarrow{p} 0$. Note that the use of the within transformation to construct the estimator implies that $\widehat{SK}_\nu$ is not affected by skewness (or kurtosis) in μ.

Consider now the estimator for skewness in the individual component, s_μ. As in the previous case, $\widehat{SK}_\mu \xrightarrow{p} s_\mu$, as $N \to \infty$ and fixed T. In addition, under the null hypothesis of normality $H_0^{s_\mu}$, $\widehat{SK}_\mu \xrightarrow{p} 0$. Note that $\widehat{SK}_\mu$ is robust to the presence of skewness (or kurtosis) in the remainder component, s_ν, even in small panels, i.e., finite T.

2.3 Kurtosis

In order to derive similar statistics to test for kurtosis, consider the fourth power of the *between* residuals, $\overline{u}_i^4 = (\mu_i + \overline{\nu}_i)^4 = \mu_i^4 + \overline{\nu}_i^4 + 4\mu_i^3\overline{\nu}_i + 4\mu_i\overline{\nu}_i^3 + 6\mu_i^2\overline{\nu}_i^2$. Taking expectations of $\overline{u}_i^4$ we obtain

$$\mathrm{E}[\overline{u}_i^4] = \mu_4 + T^{-3}(\nu_4 + 3(T-1)\sigma_\nu^4) + 6T^{-1}\sigma_\mu^2\sigma_\nu^2.$$

Consider now the fourth power of the *within* residuals, $\widetilde{u}_{it}^4 = (u_{it} - \overline{u}_i)^4 = \nu_{it}^4 + \overline{\nu}_i^4 - 4\nu_{it}^3\overline{\nu}_i - 4\nu_{it}\overline{\nu}_i^3 + 6\nu_{it}^2\overline{\nu}_i^2$. Taking expectations of $\widetilde{u}_{it}^4$ we obtain

$$\mathrm{E}[\widetilde{u}_{it}^4] = \nu_4\left(1 - 4T^{-1} + 6T^{-2} - 3T^{-3}\right) + \sigma_\nu^4(T-1)(6T^{-2} - 12T^{-3}).$$

These expectations are linear functions of ν_4 and μ_4, suggesting, again, that tests for kurtosis of each component can be based on $\overline{u}_i^4$ and $\widetilde{u}_{it}^4$. In particular, solving the equations for ν_4 and μ_4 and noting that $\widetilde{u}_{it}^4 = (u_{it} - \overline{u}_i)^4 = u_{it}^4 + \overline{u}_i^4 - 4u_{it}^3\overline{u}_i - 4u_{it}\overline{u}_i^3 + 6u_{it}^2\overline{u}_i^2$, it follows that

$$\begin{aligned}
\nu_4 &= \frac{\mathrm{E}[\widetilde{u}_{it}^4]}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}} - \frac{(T-1)(6T^{-2} - 12T^{-3})}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}}\sigma_\nu^4 \\
&= \frac{\mathrm{E}[\overline{u_i^4}] - 4\mathrm{E}[\overline{u_i^3}\overline{u}_i] + 6\mathrm{E}[\overline{u_i^2}\overline{u}_i^2] - 3\mathrm{E}[\overline{u}_i^4]}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}} - \frac{(T-1)(6T^{-2} - 12T^{-3})}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}}\sigma_\nu^4 \\
\mu_4 &= \mathrm{E}[\overline{u}_i^4] - T^{-3}(\nu_4 + 3(T-1)\sigma_\nu^4) - 6T^{-1}\sigma_\mu^2\sigma_\nu^2 \\
&= \mathrm{E}[\overline{u}_i^4]\frac{T^3 - 4T^2 + 6T}{T^3 - 4T^2 + 6T - 3} - \frac{\mathrm{E}[\overline{u_i^4}] - 4\mathrm{E}[\overline{u_i^3}\overline{u}_i] + 6\mathrm{E}[\overline{u_i^2}\overline{u}_i^2]}{T^3 - 4T^2 + 6T - 3} \\
&- \frac{(T-1)(3T^3 - 12T^2 + 12T + 3)}{(T^3 - 4T^2 + 6T - 3)T^3}\sigma_\nu^4 - \frac{6}{T}\sigma_\mu^2\sigma_\nu^2.
\end{aligned}$$

As in the previous section, the above equations are functions of the unobserved innovations. Define $\widehat{\nu}_4$ and $\widehat{\mu}_4$ as the estimators of ν_4 and μ_4, respectively, constructed using $\widehat{u}_{it}$, the OLS residuals. Thus, we have that

$$\begin{aligned}
\widehat{\nu}_4 &= \frac{\mathbb{E}[\overline{\widehat{u}_i^4}] - 4\mathbb{E}[\overline{\widehat{u}_i^3}\overline{\widehat{u}}_i] + 6\mathbb{E}[\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i^2] - 3\mathbb{E}[\overline{\widehat{u}}_i^4]}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}} - \frac{(T-1)(6T^{-2} - 12T^{-3})}{1 - 4T^{-1} + 6T^{-2} - 3T^{-3}}\widehat{\sigma}_\nu^4 \\
\widehat{\mu}_4 &= \mathbb{E}[\overline{\widehat{u}}_i^4]\frac{T^3 - 4T^2 + 6T}{T^3 - 4T^2 + 6T - 3} - \frac{\mathbb{E}[\overline{\widehat{u}_i^4}] - 4\mathbb{E}[\overline{\widehat{u}_i^3}\overline{\widehat{u}}_i] + 6\mathbb{E}[\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i^2]}{T^3 - 4T^2 + 6T - 3} \\
&- \frac{(T-1)(3T^3 - 12T^2 + 12T + 3)}{(T^3 - 4T^2 + 6T - 3)T^3}\widehat{\sigma}_\nu^4 - \frac{6}{T}\widehat{\sigma}_\mu^2\widehat{\sigma}_\nu^2.
\end{aligned}$$

Finally, the following statistics are useful to construct tests for kurtosis.

$$\widehat{KU}_\nu = \frac{\widehat{\nu}_4}{\widehat{\sigma}_\nu^4}, \tag{4}$$

$$\widehat{KU}_\mu = \frac{\widehat{\mu}_4}{\widehat{\sigma}_\mu^4}. \tag{5}$$

It is important to notice that when testing for kurtosis (or both skewness and kurtosis) based on moment conditions, skewness might influence kurtosis. This might be an important drawback of these tests. Jones, Rosco, and Pewsey (2011) show that certain kurtosis measures, when applied to certain wide families of skew-symmetric distributions, display the attractive property of skewness-invariance. The authors provide quantile-based measures of kurtosis and their interaction with skewness-inducing transformations, identifying classes of transformations that leave kurtosis measures invariant.

We use the statistics in (4) and (5) to construct tests for kurtosis. Consider first the estimator for kurtosis in the remainder component, k_ν. Under the assumptions listed below

$\widehat{KU}_\nu \xrightarrow{p} k_\nu$, for $N \to \infty$ and fixed T. Then for the purpose of testing normality, under $H_0^{k_\nu}$, $\widehat{KU}_\nu \xrightarrow{p} 3$. Note that the statistic $\widehat{KU}_\nu$ is not affected by kurtosis (or skewness) in μ. Consistent estimators for σ_ν^4 and σ_μ^4 are presented in Appendix A1.

Finally, consider the estimator for kurtosis in the individual component, k_μ. As before, $\widehat{KU}_\mu \xrightarrow{p} k_\mu$, as $N \to \infty$ and fixed T. Thus, under the null hypothesis of normality $H_0^{k_\mu}$, $\widehat{KU}_\mu \xrightarrow{p} 3$. As in the previous case, the statistic $\widehat{KU}_\mu$ is not affected by the presence of kurtosis (or skewness) in the remainder component.

3 Asymptotic theory

In order to implement the tests in practice we first derive the asymptotic distributions of the corresponding statistics developed in the previous section. Consider the regression model described by equation (1). We will impose the following regularity conditions to derive the asymptotic properties of the statistics.

Assumption 1. *Let $\{\mu_i\}$ be a sequence of i.i.d. random variables, and $\{(\boldsymbol{x}_{it}^\top, \nu_{it})\}$ be an array of i.i.d. random vectors across i and t. In addition, $\boldsymbol{x}_{it}$, μ_i and ν_{it} are independent. Both μ_i and ν_{it} have mean zero and finite* $\mathrm{E}[\mu^8]$ *and* $\mathrm{E}[\nu^8]$.

Assumption 2. $\mathrm{E}[\boldsymbol{x}_{it}\boldsymbol{x}_{it}^\top]$ *is a finite positive definite matrix. In addition,* $\mathrm{E}[||\boldsymbol{x}_{it}||^4]$ *is finite.*

Assumptions 1 and 2 are standard in the literature. The first one imposes restrictions on the sampling scheme and on the moments of the individual-specific and the remainder error components. The need to bound high order moments relates to the fact that the limiting distribution of the proposed statistics involves the variance of skewness and kurtosis statistics, which eventually depend on the sixth and eighth moments, respectively. The positive definiteness of $\mathrm{E}[\boldsymbol{x}_{it}\boldsymbol{x}_{it}^\top]$ in second assumption is a standard identification condition. The existence of $\mathrm{E}[||\boldsymbol{x}_{it}||^4]$ is usually required for the study of variance-covariance matrix. Assumptions 1 and 2 guarantee that a multivariate version of the Lindeberg-Lévy central limit theorem for $N \to \infty$ and fixed T can be used to establish the asymptotic normality of the statistics proposed in the previous section.

We construct tests for skewness and/or kurtosis based on moment conditions, as in previous work such as Jarque and Bera (1981), Bai and Ng (2005) or Bontemps and Meddahi

(2005). It is well known in this literature, and implied by the restrictions of Assumption 1, that a drawback of tests requiring high order moments is that they are not theoretically valid for numerous distributions. We will assess the finite sample performance of the proposed tests using a series of Monte Carlo experiments, reported in Section 5.

A novel aspect of our proposed framework is the ability to identify skewness and kurtosis in the error components of a standard linear panel model, separately and jointly. We provide a general framework to develop these tests and to suggest directions in which particular deviations from the model, other than non-Gaussianity, can be considered. We will focus only on the one-way error components model and leave extensions for future research. In particular, the previous assumptions could be relaxed at the cost of modeling additional terms in the expectations and variance-covariances given below for skewness and kurtosis. Suggestions for extensions are discussed in the last section.

The next theorem derives an asymptotic representation for the sample skewness and kurtosis.

Theorem 1 *Under Assumptions 1 and 2, and for $N \to \infty$ and fixed T,*

(i) $\sqrt{N}(\widehat{SK}_\nu - s_\nu) = \left(\frac{\mathcal{B}_1}{\sigma_\nu^3} - \frac{3s_\nu \mathcal{A}_1}{2\sigma_\nu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1)$

(ii) $\sqrt{N}(\widehat{SK}_\mu - s_\mu) = \left(\frac{\mathcal{B}_2}{\sigma_\mu^3} - \frac{3s_\mu \mathcal{A}_2}{2\sigma_\mu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1)$

(iii) $\sqrt{N}(\widehat{KU}_\nu - k_\nu) = \left(\frac{\mathcal{C}_1}{\sigma_\nu^4} - \frac{2k_\nu \mathcal{A}_1}{\sigma_\nu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1)$

(iv) $\sqrt{N}(\widehat{KU}_\mu - k_\mu) = \left(\frac{\mathcal{C}_2}{\sigma_\mu^4} - \frac{2k_\mu \mathcal{A}_2}{\sigma_\mu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1)$,

where $\mathbb{A}_i = (\overline{u}_i, \overline{u_i^2} - \mathrm{E}[\overline{u_i^2}], \overline{u}_i^2 - \mathrm{E}[\overline{u}_i^2], \overline{u_i^3} - \mathrm{E}[\overline{u_i^3}], \overline{u}_i^3 - \mathrm{E}[\overline{u}_i^3], \overline{u_i^4} - \mathrm{E}[\overline{u_i^4}], \overline{u}_i^4 - \mathrm{E}[\overline{u}_i^4], \overline{u_i^2}\overline{u}_i - \mathrm{E}[\overline{u_i^2}\overline{u}_i], \overline{u_i^3}\overline{u}_i - \mathrm{E}[\overline{u_i^3}\overline{u}_i], \overline{u_i^2}\overline{u}_i^2 - \mathrm{E}[\overline{u_i^2}\overline{u}_i^2])^\top$, $\mathcal{A}_1 = \big(0, \frac{1}{1-T^{-1}}, -\frac{1}{1-T^{-1}}, 0, 0, 0, 0, 0, 0, 0\big)$, $\mathcal{A}_2 = \big(0, -\frac{1}{T-1}, \frac{T}{T-1}, 0, 0, 0, 0, 0, 0, 0\big)$, $\mathcal{B}_1 = (0, 0, 0, 1, 2, 0, 0, -3, 0, 0)/(1 - 3T^{-1} + 2T^{-2})$, $\mathcal{B}_2 = \Big(-3\mathrm{E}[\overline{u}_i^2], 0, 0, \frac{-1}{T^2-3T+2}, \frac{T^2-3T}{T^2-3T+2}, 0, 0, \frac{3}{T^2-3T+2}, 0, 0\Big)$, $\mathcal{C}_1 = (0, (24T^{-2} - 12T^{-1})\sigma_\nu^2, (12T^{-1} - 24T^{-2})\sigma_\nu^2, 0, 0, 1, -3, 0, -4, 6)/(1-4T^{-1}+6T^{-2}-3T^{-3})$, and $\mathcal{C}_2 = \Big(-4\mathrm{E}[\overline{u}_i^3], -\frac{(3T^3-12T^2+12T+3)}{(T^3-4T^2+6T-3)T^2} - \frac{6\sigma_\mu^2}{T-1} + \frac{6\sigma_\nu^2}{T(T-1)}, \frac{(3T^3-12T^2+12T+3)}{(T^3-4T^2+6T-3)T^2} + \frac{6\sigma_\mu^2}{T-1} - \frac{6\sigma_\nu^2}{T-1}, 0, 0, \frac{-1}{T^3-4T^2+6T-3}, \frac{T^3-4T^2+6T}{T^3-4T^2+6T-3}, 0, \frac{4}{T^3-4T^2+6T-3}, \frac{-6}{T^3-4T^2+6T-3}\Big)$.

Proof. See Appendix A2. ■

Based on the previous asymptotic representations, the next result provides the asymptotic distributions of the statistics of interest.

Theorem 2 *Under Assumptions 1 and 2, and for $N \to \infty$ and fixed T,*

(i) $\sqrt{N}(\widehat{SK}_\nu - s_\nu)/\sqrt{\Omega^s_\nu} \xrightarrow{d} N(0,1)$, *where* $\Omega^s_\nu = NVar(\widehat{SK}_\nu)$;
(ii) $\sqrt{N}(\widehat{SK}_\mu - s_\mu)/\sqrt{\Omega^s_\mu} \xrightarrow{d} N(0,1)$, *where* $\Omega^s_\mu = NVar(\widehat{SK}_\mu)$;
(iii) $\sqrt{N}(\widehat{KU}_\nu - k_\nu)/\sqrt{\Omega^k_\nu} \xrightarrow{d} N(0,1)$, *where* $\Omega^k_\nu = NVar(\widehat{KU}_\nu)$;
(iv) $\sqrt{N}(\widehat{KU}_\mu - k_\mu)/\sqrt{\Omega^k_\mu} \xrightarrow{d} N(0,1)$, *where* $\Omega^k_\mu = NVar(\widehat{KU}_\mu)$.

Proof. See Appendix A3. ■

Theorem 2 shows that the distributions of the statistics $\widehat{SK}_\nu$, $\widehat{SK}_\mu$, $\widehat{KU}_\nu$ and $\widehat{KU}_\mu$, after centralization and standardization, are asymptotically standard normal. This result indicates that tests for skewness and/or kurtosis are simple to implement. For example, to test $H_0^{s_\mu}: s_\mu = 0$ against $H_1^{s_\mu}: s_\mu \neq 0$, a simple t-test can be used by standardizing the test statistic $\widehat{SK}_\nu$ by the square root of a consistent estimate of its variance, and imposing the null hypothesis

$$T_{s_\mu} = \widehat{SK}_\nu/\sqrt{Var(\widehat{SK}_\nu)} \xrightarrow{d} N(0,1).$$

The critical values for this two-sided test are standard. For the case of testing $H_0^{k_\nu}: k_\nu/\sigma^4_\nu = 3$ the procedure is analogous

$$T_{k_\mu} = (\widehat{KU}_\nu - 3)/\sqrt{Var(\widehat{KU}_\nu)} \xrightarrow{d} N(0,1).$$

The next section discusses how to implement the tests in practice.

4 Implementation

In this section, we develop a bootstrap procedure to estimate the variances of the skewness and kurtosis test statistics in practice. The implementation of the tests based on the results of Theorem 2 require a consistent estimator of the asymptotic variances of the corresponding statistics. Even though these variance-covariance matrices exist and are finite by the boundedness conditions in Assumption 1, they depend on the high order moments of ν and μ. Therefore their derivation requires a cumbersome calculation. Instead, we consider the bootstrap as an alternative to estimate the corresponding asymptotic variance-covariance matrices.

Although the properties of the bootstrap are widely studied for cross-section data, it is only recently that relevant results are available for panel data. Cameron and Trivedi

(2005) discuss resampling methods for panel data when N is large but T is assumed small. Kapetanios (2008) studies the bootstrap for a linear panel data model where resampling occurs in both the cross-section and time dimensions. Goncalves (2011), allowing for both temporal and cross-sectional dependence, studies the moving block bootstrap for a linear panel data model where resampling occurs only in the time dimension.

The bootstrap method we consider is implemented by randomly drawing individuals with replacement while maintaining the time series structure unaltered. This corresponds to cross-sectional resampling in Kapetanios (2008). This method resamples $\{y_{it}, \boldsymbol{x}_{it}\}$ with replacement from the cross-section dimension with probability $1/N$, maintaining the temporal structure intact for each individual i. More specifically, let $y^* = (y_{i_1}, ..., y_{i_s}, ..., y_{i_N})$ be the vector of bootstrap samples where each element of the vector of indices $(i_1, ..., i_N)$ is obtained by drawing with replacement from $(1, ..., N)$, and each element is $y_{i_s} = (y_{i_s 1}, ..., y_{i_s T})$ for $s \in (i_1, ..., i_N)$. The same vector of indices is used to obtain $\boldsymbol{x}^*$. Therefore, each bootstrap sample is given by $\{(y^*_{it}, \boldsymbol{x}^*_{it}), i = 1, \ldots, N; t = 1, \ldots, T\}$.

Next we describe the bootstrap algorithm for estimating the variances in practice. We will focus on the implementation of the bootstrap procedure for testing the null hypothesis $H_0^{s_\nu} : s_\nu = 0$ with the statistic described in equation (2) and using part (i) in Theorem 2. The algorithms for the remaining cases are analogous. To implement the test we need to compute $\widehat{\Omega}^{s*}_\nu$, the bootstrap variance estimator. The algorithm is as follows:

1. Draw B independent bootstrap samples, $(y^*_{it}, \boldsymbol{x}^*_{it})^1, (y^*_{it}, \boldsymbol{x}^*_{it})^2, ..., (y^*_{it}, \boldsymbol{x}^*_{it})^B$ each consisting of $i = 1, \ldots, N$; $t = 1, \ldots, T$, data values as described in the previous paragraph.

2. Having obtained the resampled data, evaluate the bootstrap replication corresponding to each bootstrap sample to estimate equation (1) and compute the residuals.

3. Given the residuals for each bootstrap sample, calculate the statistic of interest as in equation (2), which is denoted by $\widehat{SK}^{*b}_\nu$, $b = 1, ..., B$.

4. Finally, obtain $\widehat{\Omega}^{s*}_\nu$ with the sample variance of $\widehat{SK}^{*b}_\nu$, $b = 1, ..., B$.

After computing this estimate, the result in part (i) of Theorem 2 can be used to construct a Wald-type test as $T_{s_\mu} = N\widehat{SK}^2_\nu/\widehat{\Omega}^{s*}_\nu$. The intuition for the validity of this test is that the limit of $Var(\widehat{SK}^{*b}_\nu)$ as B goes to infinity approximates the variance of $\widehat{SK}_\nu$. The exact same procedure is used to construct the estimator of the variances of the other statistics.

The practical implementation of the tests using the bootstrap is simple. Let $\widehat{\Omega}_{\nu}^{s*}$, $\widehat{\Omega}_{\mu}^{s*}$, $\widehat{\Omega}_{\nu}^{k*}$, and $\widehat{\Omega}_{\mu}^{k*}$ be the bootstrap estimator of the variances. Then the following Wald test statistics can be used: (i) $N(\widehat{SK}_{\nu} - s_{\nu})^2/\widehat{\Omega}_{\nu}^{s*}$; (ii) $N(\widehat{SK}_{\mu} - s_{\mu})^2/\widehat{\Omega}_{\mu}^{s*}$; (iii) $N(\widehat{KU}_{\nu} - k_{\nu})^2/\widehat{\Omega}_{\nu}^{k*}$; (iv) $N(\widehat{KU}_{\mu} - k_{\mu})^2/\widehat{\Omega}_{\mu}^{k*}$. Under the corresponding null hypotheses, the statistics (i), (ii), (iii) and (iv) have χ_1^2 asymptotic distribution. Moreover, (i)+(iii) and (ii)+(iv) have χ_2^2 asymptotic distribution.

Below, we formally prove consistency of the bootstrap estimators of the variances. To show the result, we consider the following additional condition.

Assumption 3. $\frac{(\max_i |\mu_i|)^4 \vee (\max_{i,t} |\nu_{it}|)^4}{\tau_N} \xrightarrow{a.s.} 0$, *where* τ *satisfies* $\liminf_N \tau_N > 0$ *and* $\tau_N = O(e^{N^q})$ *with a* $q \in (0, 0.5)$.

Assumption 3 is a simple sufficient condition for the consistency of the bootstrap estimators of the variance-covariance estimators and is used to guarantee the technical condition (3.28) on p. 87 of Shao and Tu (1995).

Proposition 1 *Under Assumptions 1–3, the bootstrap estimator of the variances,* Ω_{ν}^{s}, Ω_{μ}^{s}, Ω_{ν}^{k}, *and* Ω_{μ}^{k}, *are consistent.*

Proof. See Appendix A4. ■

5 Monte Carlo experiments

In this section, we use simulation experiments to assess the finite sample performance of the tests discussed in the previous sections. The baseline model for the experiments is

$$y_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + \nu_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T, \tag{6}$$

where $x_{it} \sim N(0, 1)$. The parameters β_0 and β_1 are assigned a value of 1.

We study the finite sample performance of tests for skewness and kurtosis separately, and jointly, in both the individual and remainder components. As alternative sample sizes we have considered $N \in \{100, 200, 500, 1000\}$ and $T \in \{3, 5, 10\}$. We report the proportion of rejections over 1,000 Monte Carlo replications, and all results are based on a 5% nominal size. The bootstrap implementation is based on $B = 200$ bootstrap replications of the same corresponding panel sizes.

We explore the effectiveness of the proposed tests to alternative distributional processes. Table 1 reports experiments for $\nu \sim N(0,1)$ and $\mu \sim N(0,1)$. In this case, all tests should have empirical size close to 0.05. The results for the tests using bootstrap show very good empirical size for all different tests and panel sizes. In this case, tests for skewness in both ν and μ have correct size for the smallest panel size considered ($N = 100, T = 3$). Tests for kurtosis start with oversized tests for $N = 100$ (0.07 to 0.10) but reduces to 0.05 as N increases. Moreover, tests for joint skewness and kurtosis achieve correct empirical size as N increases in a similar way to the tests for kurtosis.

[Table 1 about here]

Table 2 reports experiments for $\nu \sim t_9, \mu \sim N(0,1)$ (first set of rows) and $\nu \sim N(0,1), \mu \sim t_9$ (second set of rows). The t_9-Student distribution is symmetric but presents excess kurtosis, while the 9 degrees of freedom guarantees that all required moments are finite. In the first case, tests for kurtosis in ν should have non trivial empirical power, while tests for skewness in ν and skewness and kurtosis in μ should not. In the second case, tests for kurtosis in μ should have relevant power, while tests for skewness in μ and skewness and kurtosis in ν should not. The experiments show that this is indeed the case. For $\nu \sim t_9, \mu \sim N(0,1)$, the kurtosis test for ν has power increasing in either N or T, while the remaining tests have size close to the 0.05 nominal size. For $\nu \sim N(0,1), \mu \sim t_9$, the kurtosis test for μ has power increasing only in N, while the remaining tests have size close to the 0.05 nominal size. The results show that kurtosis in one component does not affect tests for kurtosis in the other component.

[Table 2 about here]

Tables 3 and 4 report experiments for skew normal distributions generated as in Azzalini (1985).[1] The fact that skewness affects kurtosis implies that it is difficult to separate their effects in practice. To this purpose we consider Azzalini's (1985) skew normal distribution with small skewness (as given by its shape parameter set to 1) and with minimum effect on the level of kurtosis; and next we consider a skew normal distribution with large skewness and, consequently, a large effect on kurtosis (shape parameter set to 10). Table 3 (shape set

[1] We are grateful to an anonymous referee for pointing this out. See Genton (2004) for a review of skew-elliptical distributions.

to 1) reveals that our proposed tests are able to detect departures from symmetry in each component without altering the empirical size in kurtosis. That is, tests for skewness in ν have power increasing in either N or T and tests for skewness in μ have power increasing only in N when the corresponding component follows a skew normal distribution with shape 1. Moreover, the corresponding kurtosis tests have small power (the largest rejection rate is 0.128 for $N = 1000, T = 10$ in the ν test case and 0.075 for $N = 200, T = 10$ in the μ test case). Table 4 shows that a shape parameter of 10 affects the tests for both skewness and kurtosis. Overall, the results show that skewness (and kurtosis) in one component does not affect the tests for the other component.

[Table 3 about here]

[Table 4 about here]

Tables 5 and 6 reports experiments for skew t_9-Student distributions generated as in Azzalini and Capitanio (2003) with a shape parameter of 1 and 10.[2] In this case we consider skewness and kurtosis together. The simulation results show that the developed tests are responsive to both deviations in skewness and kurtosis, and that deviations in one component does not affect the empirical size in the other component.

[Table 5 about here]

[Table 6 about here]

6 Empirical application: investment models

As an empirical illustration we apply the developed tests to the Fazzari, Hubbard and Petersen (1988) investment equation model, where firm investment is regressed on a proxy for investment demand (Tobin's q) and its cash flow; a widely used model in the corporate investment literature. Following Fazzari, Hubbard, and Petersen (1988), investment–cash-flow sensitivities became a standard metric in the literature that examines the impact of financing imperfections on corporate investment. These empirical sensitivities are also used

[2] We are grateful to an anonymous referee for pointing this out. Although not reported we also produced experiments with χ_1^2 (with both skewness and kurtosis) distributions with similar results.

to draw inferences about efficiency in internal capital markets, the effect of agency on corporate spending, the role of business groups in capital allocation, and the effect of managerial characteristics on corporate policies.

Tobin's q is the ratio of the market valuation of a firm and the replacement value of its assets. Firms with a high value of q are considered attractive as investment opportunities, whereas a low value of q indicates the opposite. Investment theory is also interested in the effect of cash flow, as the theory predicts that financially constrained firms are more likely to rely on internal funds to finance investment (see Almeida, Campello and Galvao (2010), for a discussion). The baseline model in the literature is

$$I_{it}/K_{it} = \alpha + \beta q_{it-1} + \gamma CF_{it-1}/K_{it-1} + \mu_i + \nu_{it},$$

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

We check for skewness and kurtosis in both μ and ν 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 a very disappointing performance of the q theory of investment, 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 μ determines that a few firms either invest or disinvest considerably more than the rest, while kurtosis in μ determine that a few firms locate at both sides of the investment line, that is, some invest a large amount while others disinvest large amounts too. Skewness and/or kurtosis in ν show 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 (SICs 2000 to 3999) over the 2000 to 2005 period with data available from COMPUSTAT's P/S/T, full coverage. Only firms with observations in every year are used, in order 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 only consider firms that report information in each of the five years, the sample consists mainly of relatively large firms. We apply the proposed tests using $B = 200$ bootstrap repetitions.

The firm level component μ is found to be largely asymmetric (rejecting the null hypothesis $H_0^{s_\mu} : s_\mu = 0$ of symmetry with a bootstrap p-value of 0.002) but with kurtosis close to the normal value of 3 (cannot reject the null hypothesis $H_0^{k_\mu} : k_\mu = 3$ with a bootstrap p-value of 0.967). The joint test for the null hypothesis $H_0^{s_\mu \& k_\mu} : s_\mu = 0$ and $k_\mu = 3$ is also rejected with a bootstrap p-value of 0.008. The remainder component ν shows both lack of symmetry (rejecting the null hypothesis $H_0^{s_\nu} : s_\nu = 0$ of symmetry with a bootstrap p-value of 0.002) and excess kurtosis (rejecting the null hypothesis $H_0^{k_\nu} : k_\nu = 3$ with a bootstrap p-value of 0.025), and a joint test for $H_0^{s_\nu \& k_\nu} : s_\nu = 0$ and $k_\nu = 3$ with a bootstrap p-value of 0.001.

7 Conclusion and suggestions for future research

In this paper we have developed tests for skewness and/or for excess kurtosis for the one-way error components model. The tests are based on moment restrictions and are implemented after pooled OLS estimation. Besides being informative about non-Gaussian behavior, our tests can identify whether skewed errors and/or excess kurtosis arise in one or both components of the model, separately and jointly, hence providing useful information when non-normalities may affect the statistical properties of inferential strategies, and when detecting the possible presence of asymmetric or heavy tailed errors is a statistical and economic relevant question *per se*. The tests are implemented using the bootstrap method. The experiments show that empirical sizes are close to nominal ones for sample sizes similar to those used in empirical practice, and that the tests have very good power properties.

The tests are developed under the assumptions of the standard one-way error components model. Although the methodology used in this paper provides a general framework to develop these tests, deviations from the one-way error components model can be accommodated. In particular, the assumptions could be relaxed at the cost of modeling additional terms in the expectations and variance-covariances for skewness and kurtosis. For instance, if serial correlation is allowed for in the remainder component, the skewness and kurtosis statistics become more complicated, as cross-terms within each individual have to be considered. That is, in this case, the within individual cross expectations like $\mathrm{E}[\nu_{it}\nu_{ij}]$ and $\mathrm{E}[\nu_{it}\nu_{ij}\nu_{ik}]$,

$t \neq j \neq k$ will not be zero. In that case, Assumption 1 can be replaced by that in Pesaran and Tosettin (2009), and the statistics should be modified to take into account the non-zero expectations of cross terms within each individual. In addition, note that the assumptions imply that the conditional mean and variance of model (1) are well specified. In the context of the general framework specified by Wooldridge (1990, p. 18) this implies that the validity of the derived tests actually imposes more than just the hypothesis of interest by ruling out misspecification in the conditional mean and variance (i.e., heteroskedasticity in either component). If heteroskedasticity is allowed, the variance component of each error can be adjusted for differences across individuals and time-periods. This could be accommodated by the assumptions in Montes-Rojas and Sosa-Escudero (2011), that allow for heteroskedasticity and heterokurtosis. Bootstrap procedures could be used to implement these tests. In particular, the framework developed in Kapetanios (2008) and Goncalves (2011) allows for both temporal and cross-sectional dependence.

References

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

Azzalini, A. (1985). "A class of distributions which includes the normal ones," *Scandinavian Journal of Statistics*, 12, 171-178.

Azzalini, A. and Capitanio, A. (2003). "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution," *Journal of the Royal Statistical Society Series B*, 65, 367–389.

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

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

Baltagi, B.H., Song, S. H. and Koh, W., (2003). "Testing panel data regression models with spatial error correlation," *Journal of Econometrics*, 68, 133–151.

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

Bera, A. and Premaratne, G., (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 Mátyás, L., (1996). "Robustness of tests for error component models to non-normality," *Economics Letters*, 51, 161–167.

Bontemps, C. and Meddahi, N., (2005). "Testing Normality: A GMM Approach," *Journal of Econometrics*, 124, 149–186.

Cameron, A. C., and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press, Cambridge, Massachusetts.

Dufour, J-M., Khalaf, L. and Beaulieu, M-C., (2003). "Exact skewness-kurtosis tests for multivariate normality and goodness-of-fit in multivariate regressions with application to asset pricing models," *Oxford Bulletin of Economics and Statistics*, 65, 891– 906.

Ergun, A.T. and Jun, J., (2010). "Conditional skewness, kurtosis, and density specification testing: Moment-based versus nonparametric tests," *Studies in Nonlinear Dynamics & Econometrics*, 14, Article 5.

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

Evans, M., (1992). "Robustness of size of tests of autocorrelation and heteroskedasticity to nonnormality," *Journal of Econometrics*, 51, 7–24.

Fazzari, S., R. G. Hubbard, and B. Petersen (1988). "Financing constraints and corporate investment," *Brooking Papers on Economic Activity,*, 1, 141–195.

Genton, G. (2004). *Skew-Elliptical Distributions and Their Applications. A Journey Beyond Normality*, Chapman & Hall/CRC, London.

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

Goncalves, S. (2011). "The moving blocks bootstrap for panel linear regression models with individual fixed effects," *Econometric Theory*, 27, 1048–1082.

Henze, N. (1994). "On Mardias kurtosis test for multivariate normality," *Communications in Statistics, Part A - Theory and Methods*, 23, pp. 1031–1045.

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

Jarque, C.M. and Bera, A.K., (1981). "Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence,". *Economics Letters*, 7, 313–318.

Jones, M. C., Rosco, J. F., and Pewsey, A. (2011). "Skewness-invariant measures of kurtosis,". *The American Statistician*, 65, 89–95.

Kapetanios, G. (2008). "A bootstrap procedure for panel datasets with many cross-sectional units," *Econometrics Journal*, 11, 377–395.

Lutkepohl, H. and Theilen, B. (1991). "Measures of multivariate skewness and kurtosis for tests of nonnormality," *Statistical Papers*, 32, 179–193.

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

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

Pesaran, M. H. and Tosetti, E., (2009). "Large panels with spatial correlation and common factors," *Working Paper.*

Premaratne, G. and Bera, A., (2005), "A test for symmetry with leptokurtic financial data," *Journal of Financial Econometrics*, 3, 169–187.

Shao, J. and Tu, D., (1995). *The Jackknife and Bootstrap*, Springer-Verlag, New York, New York.

Wooldridge, J.M., (1990). "A unified approach to robust, regression-based specification tests," *Econometric Theory*, 6, 17–43.

Appendix A

A1. Variance

We present estimators of the variance of μ and ν. First, we derive some equalities based on which we estimate the variance of μ and ν. Note that $\overline{u}_i^2 = (\mu_i + \overline{\nu}_i)^2 = \mu_i^2 + \overline{\nu}_i^2 + 2\mu_i\overline{\nu}_i$. Taking expectation, we have

$$\mathrm{E}[\overline{u}_i^2] = \sigma_\mu^2 + T^{-1}\sigma_\nu^2.$$

The square of the *within* residuals is given by $\widetilde{u}_{it}^2 = (u_{it} - \overline{u}_i)^2 = \nu_{it}^2 + \overline{\nu}_i^2 - 2\nu_{it}\overline{\nu}_i$, and after taking expectations

$$\mathrm{E}[\widetilde{u}_{it}^2] = \sigma_\nu^2 + T^{-1}\sigma_\nu^2 - 2T^{-1}\sigma_\nu^2 = \sigma_\nu^2\left(1 - T^{-1}\right).$$

Solving these equations for σ_μ^2 and σ_ν^2, we obtain

$$\begin{aligned}
\sigma_\nu^2 &= \frac{1}{1-T^{-1}}\mathrm{E}[\overline{\widetilde{u}_i^2}] = \frac{1}{1-T^{-1}}\mathrm{E}[\overline{u_i^2}] - \frac{1}{1-T^{-1}}\mathrm{E}[\overline{u}_i^2],\\
\sigma_\mu^2 &= \mathrm{E}[\overline{u}_i^2] - \frac{1}{T-1}\mathrm{E}[\overline{\widetilde{u}_i^2}] = \mathrm{E}[\overline{u}_i^2] - \frac{1}{T-1}\mathrm{E}[\overline{u_i^2} - \overline{u}_i^2] = \frac{T}{T-1}\mathrm{E}[\overline{u}_i^2] - \frac{1}{T-1}\mathrm{E}[\overline{u_i^2}].
\end{aligned}$$

Next, given this derivation, we consider the following statistics that are the sample counterpart of the above expressions, respectively, by replacing the theoretical with the sample expectations and using $\widehat{u}_{it}$, the OLS residuals

$$\begin{aligned}
\widehat{\sigma}_\nu^2 &= \frac{1}{1-T^{-1}}\mathbb{E}[\overline{\widehat{u}_i^2}] - \frac{1}{1-T^{-1}}\mathbb{E}[\overline{\widehat{u}}_i^2] = \frac{1}{1-T^{-1}}\mathbb{E}[\overline{u_i^2}] - \frac{1}{1-T^{-1}}\mathbb{E}[\overline{u}_i^2] + o_p(N^{-1/2}),\\
\widehat{\sigma}_\mu^2 &= \frac{T}{T-1}\mathbb{E}[\overline{\widehat{u}}_i^2] - \frac{1}{T-1}\mathbb{E}[\overline{\widehat{u}_i^2}] = \frac{T}{T-1}\mathbb{E}[\overline{u}_i^2] - \frac{1}{T-1}\mathbb{E}[\overline{u_i^2}] + o_p(N^{-1/2}).
\end{aligned}$$

The second equalities of the two lines above follows from Lemmas 1 and 2 (in Appendix B) with $j = 2$. Therefore,

$$\begin{aligned}
\sqrt{N}(\widehat{\sigma}_\nu^2 - \sigma_\nu^2) &= \mathcal{A}_1 \frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbb{A}_i + o_p(1),\\
\sqrt{N}(\widehat{\sigma}_\mu^2 - \sigma_\mu^2) &= \mathcal{A}_2 \frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbb{A}_i + o_p(1),
\end{aligned}$$

and by delta method and Slutsky's lemma, respectively,

$$\sqrt{N}(\widehat{\sigma}_\nu^4 - \sigma_\nu^4) = 2\sigma_\nu^2 \mathcal{A}_1 \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1),$$

$$\sqrt{N}(\widehat{\sigma}_\mu^2 \widehat{\sigma}_\nu^2 - \sigma_\mu^2 \sigma_\nu^2) = (\sigma_\mu^2 \mathcal{A}_1 + \sigma_\nu^2 \mathcal{A}_2) \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1).$$

A2. Proof of Theorem 1

A2.1 Proof of skewness results in (i) and (ii)

Note that

$$\sqrt{N}(\widehat{SK_\nu} - s_\nu) = \sqrt{N}\left(\frac{\widehat{\nu}_3}{\widehat{\sigma}_\nu^3} - \frac{\nu_3}{\sigma_\nu^3}\right) = \sqrt{N}\left[\frac{\widehat{\nu}_3 - \nu_3}{\widehat{\sigma}_\nu^3} - \frac{3 s_\nu \sigma_\nu (\widehat{\sigma}_\nu^2 - \sigma_\nu^2)}{2\widehat{\sigma}_\nu^3}\right] + o_p(1),$$

$$\sqrt{N}(\widehat{SK_\mu} - s_\mu) = \sqrt{N}\left(\frac{\widehat{\mu}_3}{\widehat{\sigma}_\mu^3} - \frac{\mu_3}{\sigma_\mu^3}\right) = \sqrt{N}\left[\frac{\widehat{\mu}_3 - \mu_3}{\widehat{\sigma}_\mu^3} - \frac{3 s_\mu \sigma_\mu (\widehat{\sigma}_\mu^2 - \sigma_\mu^2)}{2\widehat{\sigma}_\mu^3}\right] + o_p(1).$$

Then we only need to find asymptotic linear representations of $\sqrt{N}(\widehat{\nu}_3 - \nu_3)$ and $\sqrt{N}(\widehat{\mu}_3 - \mu_3)$, since those of $\sqrt{N}(\widehat{\sigma}_\nu^2 - \sigma_\nu^2)$ and $\sqrt{N}(\widehat{\sigma}_\mu^2 - \sigma_\mu^2)$ are obtained in Appendix A1. Using the conclusions of Lemmas 1 and 2 (in Appendix B) for $j = 3$ and the first equality of Lemma 3 (in Appendix B), we have

$$\begin{aligned}
&\sqrt{N}(\widehat{\nu}_3 - \nu_3) = \sqrt{N}\left(\frac{1}{1 - 3T^{-1} + 2T^{-2}} \mathbb{E}[\overline{\widehat{u}_i^3} - 3\overline{\widehat{u}}_i \overline{\widehat{u}_i^2} + 2\overline{\widehat{u}}_i^3] - \nu_3\right) \\
=& \frac{1}{1 - 3T^{-1} + 2T^{-2}} \left[\left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u_i^3} - \sqrt{N}\mathrm{E}[\overline{u_i^3}] - 3\mathrm{E}[u_{it}^2] \frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u}_i\right)\right. \\
&- 3\left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u_i^2}\,\overline{u}_i - \sqrt{N}\mathrm{E}[\overline{u_i^2}\,\overline{u}_i] - \frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u}_i \mathrm{E}[u_{it}^2 + 2\overline{u}_i^2]\right) \\
&\left. +2\left(\frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u}_i^3 - \sqrt{N}\mathrm{E}[\overline{u}_i^3] - 3\mathrm{E}[\overline{u}_i^2] \frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u}_i\right)\right] + o_p(1) \\
=& \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u_i^3} - \sqrt{N}\mathrm{E}[\overline{u_i^3}] - 3\frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u_i^2}\,\overline{u}_i + 3\sqrt{N}\mathrm{E}[\overline{u_i^2}\,\overline{u}_i] + 2\frac{1}{\sqrt{N}} \sum_{i=1}^N \overline{u}_i^3 - 2\sqrt{N}\mathrm{E}[\overline{u}_i^3]}{1 - 3T^{-1} + 2T^{-2}} + o_p(1) \\
=& \mathcal{B}_1 \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{A}_i + o_p(1).
\end{aligned}$$

$$
\begin{aligned}
&\sqrt{N}(\widehat{\mu}_3-\mu_3)=\sqrt{N}\left(\frac{T^2-3T}{T^2-3T+2}\mathbb{E}[\overline{\widehat{u}}_i^3]-\frac{1}{T^2-3T+2}\mathbb{E}[\overline{\widehat{u}_i^3}-3\overline{\widehat{u}}_i\overline{\widehat{u}_i^2}]-\mu_3\right)\\
=&\frac{T^2-3T}{T^2-3T+2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i^3-\sqrt{N}\mathrm{E}[\overline{u}_i^3]-3\mathrm{E}[\overline{u}_i^2]\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i\right)\\
&-\frac{1}{T^2-3T+2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^3}-\sqrt{N}\mathrm{E}[\overline{u_i^3}]-3\mathrm{E}[u_{it}^2]\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i\right)\\
&+3\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i-\sqrt{N}\mathrm{E}[\overline{u_i^2}\overline{u}_i]-\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i\mathrm{E}[u_{it}^2+2\overline{u}_i^2]\right)\div(T^2-3T+2)+o_p(1)\\
=&\frac{T^2-3T}{T^2-3T+2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i^3-\sqrt{N}\mathrm{E}[\overline{u}_i^3]\right)-\frac{1}{T^2-3T+2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^3}-\sqrt{N}\mathrm{E}[\overline{u_i^3}]\right)\\
&+\frac{3}{T^2-3T+2}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i-\sqrt{N}\mathrm{E}[\overline{u_i^2}\overline{u}_i]\right)-3\mathrm{E}[\overline{u}_i^2]\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i+o_p(1)\\
=&\mathcal{B}_2\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbb{A}_i+o_p(1).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\sqrt{N}(\widehat{SK}_\nu-s_\nu)&=\left(\frac{\mathcal{B}_1}{\sigma_\nu^3}-\frac{3s_\nu\mathcal{A}_1}{2\sigma_\nu^2}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbb{A}_i+o_p(1)\\
\sqrt{N}(\widehat{SK}_\mu-s_\mu)&=\left(\frac{\mathcal{B}_2}{\sigma_\mu^3}-\frac{3s_\mu\mathcal{A}_2}{2\sigma_\mu^2}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbb{A}_i+o_p(1).
\end{aligned}
$$

A2.2 Proof of kurtosis results in (iii) and (iv)

Note that

$$
\begin{aligned}
\sqrt{N}(\widehat{KU}_\nu-k_\nu)&=\sqrt{N}\left(\frac{\widehat{\nu}_4}{\widehat{\sigma}_\nu^4}-\frac{\nu_4}{\sigma_\nu^4}\right)=\sqrt{N}\left[\frac{\widehat{\nu}_4-\nu_4}{\widehat{\sigma}_\nu^4}-\frac{2k_\nu\sigma_\nu^2(\widehat{\sigma}_\nu^2-\sigma_\nu^2)}{\widehat{\sigma}_\nu^4}\right]+o_p(1),\\
\sqrt{N}(\widehat{KU}_\mu-k_\mu)&=\sqrt{N}\left(\frac{\widehat{\mu}_4}{\widehat{\sigma}_\mu^4}-\frac{\mu_4}{\sigma_\mu^4}\right)=\sqrt{N}\left[\frac{\widehat{\mu}_4-\mu_4}{\widehat{\sigma}_\mu^4}-\frac{2k_\mu\sigma_\mu^2(\widehat{\sigma}_\mu^2-\sigma_\mu^2)}{\widehat{\sigma}_\mu^4}\right]+o_p(1).
\end{aligned}
$$

Then we only need to find asymptotic linear representations of $\sqrt{N}(\widehat{\nu}_4-\nu_4)$ and $\sqrt{N}(\widehat{\mu}_4-\mu_4)$, since those of $\sqrt{N}(\widehat{\sigma}_\nu^2-\sigma_\nu^2)$ and $\sqrt{N}(\widehat{\sigma}_\mu^2-\sigma_\mu^2)$ are obtained in Appendix A1. Using the conclusions from Lemmas 1 and 2 with $j=4$ and the second and third equalities from

Lemma 3 (in Appendix B),

$$
\begin{aligned}
&\sqrt{N}(\widehat{\nu}_4-\nu_4)=\sqrt{N}\left(\frac{\mathbb{E}[\overline{\widehat{u}_i^4}]-4\mathbb{E}[\overline{\widehat{u}_i^3}\overline{\widehat{u}}_i]+6\mathbb{E}[\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i^2]-3\mathbb{E}[\overline{\widehat{u}}_i^4]}{1-4T^{-1}+6T^{-2}-3T^{-3}}-\frac{(T-1)(6T^{-2}-12T^{-3})}{1-4T^{-1}+6T^{-2}-3T^{-3}}\widehat{\sigma}_\nu^4-\nu_4\right)\\
=&\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\frac{1}{T}\sum_{t=1}^T u_{it}^4-\sqrt{N}\mathrm{E}[\overline{u_i^4}]-4\mathrm{E}[u_{it}^3]\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\right)\right.\\
&-4\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u_i^3}\overline{u}_i-\sqrt{N}\mathrm{E}[\overline{u_i^3}\overline{u}_i]-\mathrm{E}[3\overline{u_i^2}\overline{u}_i+u_{it}^3]\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\right)\\
&+6\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u_i^2}\overline{u}_i^2-\sqrt{N}\mathrm{E}[\overline{u_i^2}\overline{u}_i^2]-2\mathrm{E}[\overline{u}_i^3+\overline{u_i^2}\overline{u}_i]\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\right)\\
&\left.-3\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i^4-\sqrt{N}\mathrm{E}[\overline{u}_i^4]-4\mathrm{E}[\overline{u}_i^3]\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\right)\right]\div(1-4T^{-1}+6T^{-2}-3T^{-3})\\
&-\frac{(T-1)(6T^{-2}-12T^{-3})}{1-4T^{-1}+6T^{-2}-3T^{-3}}2\sigma_\nu^2\mathcal{A}_1\frac{1}{\sqrt{N}}\sum_{i=1}^N\mathbb{A}_i+o_p(1)\\
=&\mathcal{C}_1\frac{1}{\sqrt{N}}\sum_{i=1}^N\mathbb{A}_i+o_p(1),
\end{aligned}
$$

$$
\begin{aligned}
&\sqrt{N}(\widehat{\mu}_4-\mu_4)=\sqrt{N}\left(\mathbb{E}[\overline{\widehat{u}}_i^4]\frac{T^3-4T^2+6T}{T^3-4T^2+6T-3}-\frac{\mathbb{E}[\overline{\widehat{u}_i^4}]-4\mathbb{E}[\overline{\widehat{u}_i^3}\overline{\widehat{u}}_i]+6\mathbb{E}[\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i^2]}{T^3-4T^2+6T-3}\right.\\
&\left.-\frac{(T-1)(3T^3-12T^2+12T+3)}{(T^3-4T^2+6T-3)T^3}\widehat{\sigma}_\nu^4-\frac{6}{T}\widehat{\sigma}_\mu^2\widehat{\sigma}_\nu^2-\mu_4\right)\\
=&\frac{T^3-4T^2+6T}{T^3-4T^2+6T-3}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i^4-\sqrt{N}\mathrm{E}[\overline{u}_i^4]-4\mathrm{E}[\overline{u}_i^3]\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\right)\\
&-\frac{1}{T^3-4T^2+6T-3}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u_i^4}-\sqrt{N}\mathrm{E}[\overline{u_i^4}]-4\mathrm{E}[u_{it}^3]\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\right)\\
&+\frac{4}{T^3-4T^2+6T-3}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u_i^3}\overline{u}_i-\sqrt{N}\mathrm{E}[\overline{u_i^3}\overline{u}_i]-\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\mathrm{E}[3\overline{u_i^2}\overline{u}_i+u_{it}^3]\right)\\
&-\frac{6}{T^3-4T^2+6T-3}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u_i^2}\overline{u}_i^2-\sqrt{N}\mathrm{E}[\overline{u_i^2}\overline{u}_i^2]-2\frac{1}{\sqrt{N}}\sum_{i=1}^N\overline{u}_i\mathrm{E}[\overline{u}_i^3+\overline{u_i^2}\overline{u}_i]\right)\\
&-\frac{(T-1)(3T^3-12T^2+12T+3)}{(T^3-4T^2+6T-3)T^3}\left(2\sigma_\nu^2\mathcal{A}_1\frac{1}{\sqrt{N}}\sum_{i=1}^N\mathbb{A}_i\right)-\frac{6}{T}\left((\sigma_\mu^2\mathcal{A}_1+\sigma_\nu^2\mathcal{A}_2)\frac{1}{\sqrt{N}}\sum_{i=1}^N\mathbb{A}_i\right)+o_p(1)\\
=&\mathcal{C}_2\frac{1}{\sqrt{N}}\sum_{i=1}^N\mathbb{A}_i+o_p(1).
\end{aligned}
$$

Therefore,

$$\sqrt{N}(\widehat{KU_\nu} - k_\nu) = \left(\frac{\mathcal{C}_1}{\sigma_\nu^4} - \frac{2k_\nu \mathcal{A}_1}{\sigma_\nu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{A}_i + o_p(1)$$
$$\sqrt{N}(\widehat{KU_\mu} - k_\mu) = \left(\frac{\mathcal{C}_2}{\sigma_\mu^4} - \frac{2k_\mu \mathcal{A}_2}{\sigma_\mu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{A}_i + o_p(1).$$

A3. Proof of Theorem 2

We prove the conclusion (i) only because the conclusions of the other three are similar. By Lindeberg-Lévy central limit theorem and Assumption 1, $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{A}_i$ converges to a normal distribution. Then, using the continuous mapping theorem, $\left(\frac{\mathcal{B}_1}{\sigma_\nu^3} - \frac{3s_\nu \mathcal{A}_1}{2\sigma_\nu^2}\right) \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbb{A}_i$ converges to a normal distribution with mean zero and variance-covariance matrix denoted by Ω_ν^s. Applying the continuous mapping theorem again, we obtain (i).

A4. Proof of Proposition 1

We will prove only the consistency of the bootstrap estimator of Ω_ν^s as the the consistency of the other estimators are similar. We apply Theorem 3.8 of Shao and Tu (1995) as the statistic is an average $\mathbb{E}\left[\left(\frac{\mathcal{B}_1}{\sigma_\nu^3} - \frac{3s_\nu \mathcal{A}_1}{2\sigma_\nu^2}\right) \mathbb{B}_i\right]$, where $\mathbb{B}_i = (\overline{u}_i, \overline{u_i^2}, \overline{u}_i^2, \overline{u_i^3}, \overline{u}_i^3, \overline{u_i^4}, \overline{u}_i^4, \overline{u_i^2}\overline{u}_i, \overline{u_i^3}\overline{u}_i, \overline{u_i^2}\overline{u}_i^2)^\top$. We focus on verifying condition (3.28). It suffices to show that

$$\frac{\max_{1\le i\le N} \left| \left(\frac{\mathcal{B}_1}{\sigma_\nu^3} - \frac{3s_\nu \mathcal{A}_1}{2\sigma_\nu^2}\right) \mathbb{B}_i - \mathbb{E}\left[\left(\frac{\mathcal{B}_1}{\sigma_\nu^3} - \frac{3s_\nu \mathcal{A}_1}{2\sigma_\nu^2}\right) \mathbb{B}_i\right] \right|}{\tau_N} \xrightarrow{a.s.} 0,$$

which is implied by Assumption 3.

Appendix B

Auxiliary Lemmas

In this appendix we list some results that are useful for the proofs of the theorems.

Lemma 1 *Under Assumptions 1 and 2, for j =2, 3, and 4, we have*

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \overline{\widehat{u}_i^j} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \overline{u_i^j} - j\mathrm{E}[u_{it}^{j-1}] \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \overline{u}_i + o_p(1).$$

Proof. The proof is complete if the following equalities hold.

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\widehat{u}_{it}^{j} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}(u_{it}-\overline{u})^{j}+o_p(1),$$

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}(u_{it}-\overline{u})^{j} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}u_{it}^{j}-j\mathrm{E}[u_{it}^{j-1}]\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}u_{it}+o_p(1).$$

The proofs of the first and second equalities are modifications of the proofs of Theorem 5 and Lemma A.1., respectively, of Bai and Ng (2005) to accommodate the one-way error components panel data model. ■

Lemma 2 *Under Assumptions 1 and 2, for j =2, 3, and 4, we have*

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}}_{i}^{j} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_{i}^{j}-j\mathrm{E}[\overline{u}_{i}^{j-1}]\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_{i}+o_p(1).$$

Proof. Noting that $\widehat{u}_{it}=(u_{it}-\overline{u})-(\boldsymbol{x}_{it}-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)$, we have $\overline{\widehat{u}}_{i}=(\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)$. Also, we will use the fact that $\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0=O_p(N^{-1/2})$ and that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^{a}\left((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\right)^{b}=O_p(N^{-1/2})$$

where $a\geq 0$ and $b>0$ are integers. In fact, whenever $a\geq 0$ and $b\geq 2$, if $\mathrm{E}[(\overline{u}_i-\overline{u})^{a}||\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}}||^{b}]$ exists, the expression equals

$$O_p(N^{1/2})O_p(N^{-b/2})=O_p(N^{(1-b)/2})=O_p(N^{-1/2}).$$

When $a\geq 0$ and $b=1$,

$$\begin{aligned}&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^{a}(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\\ =&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}((\overline{u}_i-\overline{u})^{a}-\mathrm{E}[(\overline{u}_i-\overline{u})^{a}])(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)=O_p(1)O_p(N^{-1/2}).\end{aligned}$$

For $j=2$,

$$\begin{aligned}&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}}_{i}^{2}=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left((\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\right)^{2}\\ =&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^{2}+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\right)^{2}-2\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})\left((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^{\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\right)\\ =&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\overline{u}_i^{2}+\overline{u}^{2}-2\overline{u}_i\overline{u}\right)+O_p(N^{1/2})O_p(N^{-1})-O_p(1)O_p(N^{-1/2})=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i^{2}+o_p(1).\end{aligned}$$

For $j=3$,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}}_i^3=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Big((\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^3$$

$$=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^3-\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^3-3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^2\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)$$

$$+3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^2$$

$$=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\overline{u}_i^3-\overline{u}^3-3\overline{u}_i^2\overline{u}+3\overline{u}_i\overline{u}^2\right)-O_p(N^{1/2})O_p(N^{-3/2})-O_p(1)O_p(N^{-1/2})+O_p(1)O_p(N^{-1})$$

$$=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i^3-3(\sigma_\mu^2+T^{-1}\sigma_\nu^2)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i+o_p(1).$$

For $j=4$,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}}_i^4=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Big((\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^4$$

$$=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^4+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^4-4\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^3\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)$$

$$-4\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^3+6\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u}_i-\overline{u})^2\Big((\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\Big)^2$$

$$=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\overline{u}_i^4+\overline{u}^4-4\overline{u}_i^3\overline{u}-4\overline{u}_i\overline{u}^3+6\overline{u}_i^2\overline{u}^2\right)+O_p(N^{-1/2})$$

$$=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i^4-4(\mu_3+T^{-2}\nu_3)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u}_i+o_p(1).$$

■

Lemma 3 *Under Assumptions 1 and 2, the following equalities hold.*

$$(i)\quad \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i-\sqrt{N}\overline{u}\mathrm{E}[u_{it}^2+2\overline{u}_i^2]+o_p(1)$$

$$(ii)\quad \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}_i^3}\overline{\widehat{u}}_i=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^3}\overline{u}_i-\sqrt{N}\overline{u}\mathrm{E}[3\overline{u_i^2}\overline{u}_i+u_{it}^3]+o_p(1)$$

$$(iii)\quad \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i^2=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i^2-2\sqrt{N}\overline{u}\mathrm{E}[\overline{u}_i^3+\overline{u_i^2}\overline{u}_i]+o_p(1).$$

Proof. To derive the first equality, we conduct the following calculations.

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}_i^2}\,\overline{\widehat{u}}_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{[(u_i-\overline{u})-(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^2}\times\overline{[(u_i-\overline{u})-(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]}\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\overline{(u_i-\overline{u})^2}-2\overline{(u_i-\overline{u})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+\overline{[(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^2}][(\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[(\overline{u_i^2}-2\overline{u_i}\,\overline{u}+\overline{u}^2)-2\overline{(u_i-\overline{u})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+O_p(N^{-1})][(\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u_i^2}-2\overline{u_i}\,\overline{u}+\overline{u}^2)(\overline{u}_i-\overline{u})-2(\overline{u}_i-\overline{u})\overline{(u_i-\overline{u})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\\
&-(\overline{u_i^2}-2\overline{u_i}\,\overline{u}+\overline{u}^2)(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+o_p(1)\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i-\sqrt{N}\overline{u}\mathrm{E}[u_{it}^2+2\overline{u}_i^2]-2\sqrt{N}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)^\top(\mathrm{E}[\overline{u}_i\overline{u_i\boldsymbol{x}_i}]-\mathrm{E}[\overline{u}_i^2\overline{\boldsymbol{x}}])+o_p(1)\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i-\sqrt{N}\overline{u}\mathrm{E}[u_{it}^2+2\overline{u}_i^2]+o_p(1).
\end{aligned}
$$

Now we establish the second equality.

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}_i^3}\,\overline{\widehat{u}}_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{[(u_i-\overline{u})-(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^3}\times\overline{[(u_i-\overline{u})-(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]}\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\overline{(u_i-\overline{u})^3}-3\overline{(u_i-\overline{u})^2(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+3\overline{(u_i-\overline{u})[(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^2}\\
&-\overline{[(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^3}][(\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[(\overline{u_i^3}-3\overline{u_i^2}\,\overline{u}+3\overline{u_i}\,\overline{u}^2-\overline{u}^3)-3\overline{(u_i^2-2u_i\overline{u}+\overline{u}^2)(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+O_p(N^{-1})]\\
&\times[(\overline{u}_i-\overline{u})-(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u_i^3}-3\overline{u_i^2}\,\overline{u}+3\overline{u_i}\,\overline{u}^2-\overline{u}^3)(\overline{u}_i-\overline{u})-3(\overline{u}_i-\overline{u})\overline{(u_i^2-2u_i\overline{u}+\overline{u}^2)(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\\
&-(\overline{u_i^3}-3\overline{u_i^2}\,\overline{u}+3\overline{u_i}\,\overline{u}^2-\overline{u}^3)(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+o_p(1)\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^3}\overline{u}_i-\sqrt{N}\overline{u}\mathrm{E}[3\overline{u_i^2}\overline{u}_i+u_{it}^3]-2\sqrt{N}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)^\top(\mathrm{E}[\overline{u}_i\overline{u_i^2\boldsymbol{x}_i}]-\mathrm{E}[\overline{u}_i\overline{u_i^2}\overline{\boldsymbol{x}}])+o_p(1)\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^3}\overline{u}_i-\sqrt{N}\overline{u}\mathrm{E}[3\overline{u_i^2}\overline{u}_i+u_{it}^3]+o_p(1).
\end{aligned}
$$

Finally, we show the third equality.

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{\widehat{u}_i^2}\overline{\widehat{u}}_i^2 = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{[(u_i-\overline{u})-(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^2}\times\overline{[(u_i-\overline{u})-(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]}^2\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[\overline{(u_i-\overline{u})^2}-2\overline{(u_i-\overline{u})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+\overline{[(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^2}]\\
&\times[(\overline{u}_i-\overline{u})^2-2(\overline{u}_i-\overline{u})(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+[(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)]^2]\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[(\overline{u_i^2}-2\overline{u_i}\,\overline{u}+\overline{u}^2)-2\overline{(u_i-\overline{u})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+O_p(N^{-1})]\\
&\times[(\overline{u}_i^2-2\overline{u}_i\overline{u}+\overline{u}^2)-2(\overline{u}_i-\overline{u})(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+O_p(T^{-1})]\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\overline{u_i^2}-2\overline{u_i}\,\overline{u}+\overline{u}^2)(\overline{u}_i^2-2\overline{u}_i\overline{u}+\overline{u}^2)-(\overline{u}_i^2-2\overline{u_i}\,\overline{u}+\overline{u}^2)2\overline{(u_i-\overline{u})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})^\top}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)\\
&(\overline{u_i^2}-2\overline{u}_i\overline{u}+\overline{u}^2)2(\overline{u}_i-\overline{u})(\overline{\boldsymbol{x}}_i-\overline{\boldsymbol{x}})^\top(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)+o_p(1)\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i^2-2\sqrt{N}\overline{u}\mathrm{E}[\overline{u}_i^3+\overline{u_i^2}\overline{u}_i]+2\sqrt{N}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)^\top\mathrm{E}[\overline{u}_i^3\overline{\boldsymbol{x}}-\overline{u}_i^2\overline{u_i\boldsymbol{x}_i}]+o_p(1)\\
=&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\overline{u_i^2}\overline{u}_i^2-2\sqrt{N}\overline{u}\mathrm{E}[\overline{u}_i^3+\overline{u_i^2}\overline{u}_i]+o_p(1).
\end{aligned}
$$

■

Table 1: $\nu \sim N(0,1), \mu \sim N(0,1)$

		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.053	0.082	0.074	0.053	0.038	0.044
100	5	0.052	0.102	0.090	0.056	0.068	0.057
100	10	0.048	0.080	0.082	0.052	0.071	0.051
200	3	0.057	0.078	0.069	0.064	0.054	0.065
200	5	0.051	0.082	0.083	0.059	0.079	0.068
200	10	0.055	0.065	0.066	0.063	0.100	0.081
500	3	0.045	0.071	0.064	0.041	0.053	0.051
500	5	0.047	0.058	0.052	0.055	0.073	0.070
500	10	0.054	0.060	0.056	0.059	0.069	0.069
1000	3	0.055	0.058	0.059	0.052	0.052	0.054
1000	5	0.044	0.052	0.048	0.055	0.059	0.070
1000	10	0.062	0.053	0.065	0.042	0.075	0.065

Notes: Monte Carlo experiments based on $R = 1,000$ replications. Bootstrap implementation uses $B = 200$ bootstrap replications.

Table 2: $\nu \sim t_9, \mu \sim N(0,1)$ and $\nu \sim N(0,1), \mu \sim t_9$

$\nu \sim t_9, \mu \sim N(0,1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.032	0.051	0.049	0.045	0.060	0.046
100	5	0.051	0.211	0.137	0.050	0.067	0.054
100	10	0.043	0.665	0.497	0.046	0.076	0.064
200	3	0.048	0.130	0.095	0.043	0.066	0.050
200	5	0.048	0.502	0.337	0.048	0.083	0.073
200	10	0.048	0.903	0.806	0.049	0.067	0.059
500	3	0.043	0.511	0.340	0.056	0.049	0.058
500	5	0.047	0.902	0.805	0.052	0.072	0.067
500	10	0.044	0.984	0.955	0.049	0.074	0.074
1000	3	0.045	0.850	0.721	0.046	0.054	0.057
1000	5	0.058	0.978	0.956	0.060	0.071	0.070
1000	10	0.051	0.992	0.986	0.053	0.078	0.076
$\nu \sim N(0,1), \mu \sim t_9$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.068	0.081	0.064	0.046	0.041	0.054
100	5	0.047	0.088	0.080	0.041	0.045	0.038
100	10	0.045	0.068	0.061	0.062	0.050	0.062
200	3	0.049	0.077	0.060	0.060	0.089	0.085
200	5	0.050	0.057	0.066	0.034	0.076	0.052
200	10	0.052	0.054	0.062	0.034	0.071	0.050
500	3	0.038	0.083	0.068	0.052	0.237	0.165
500	5	0.043	0.074	0.066	0.053	0.258	0.176
500	10	0.064	0.075	0.058	0.038	0.304	0.181
1000	3	0.051	0.076	0.066	0.041	0.584	0.404
1000	5	0.052	0.060	0.064	0.046	0.579	0.404
1000	10	0.042	0.074	0.058	0.052	0.666	0.490

Notes: Monte Carlo experiments based on $R = 1,000$ replications. Bootstrap implementation uses $B = 200$ bootstrap replications.

Table 3: $\nu \sim skew-normal(shape = 1), \mu \sim N(0,1)$ and $\nu \sim N(0,1), \mu \sim skew-normal(shape = 1)$

$\nu \sim skew-normal(shape=1), \mu \sim N(0,1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.058	0.087	0.088	0.049	0.057	0.045
100	5	0.138	0.061	0.112	0.040	0.071	0.062
100	10	0.324	0.055	0.237	0.059	0.056	0.053
200	3	0.081	0.080	0.095	0.041	0.053	0.043
200	5	0.214	0.046	0.161	0.047	0.082	0.060
200	10	0.534	0.046	0.442	0.053	0.093	0.074
500	3	0.160	0.069	0.133	0.059	0.050	0.049
500	5	0.490	0.054	0.393	0.050	0.069	0.060
500	10	0.915	0.054	0.839	0.048	0.085	0.074
1000	3	0.294	0.049	0.218	0.049	0.059	0.053
1000	5	0.779	0.056	0.690	0.055	0.060	0.058
1000	10	0.997	0.128	0.993	0.052	0.076	0.065
$\nu \sim N(0,1), \mu \sim skew-normal(shape=1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.052	0.091	0.083	0.063	0.046	0.048
100	5	0.054	0.072	0.065	0.078	0.063	0.063
100	10	0.040	0.078	0.065	0.069	0.059	0.061
200	3	0.050	0.077	0.076	0.084	0.054	0.063
200	5	0.052	0.083	0.073	0.090	0.064	0.089
200	10	0.057	0.053	0.060	0.077	0.075	0.077
500	3	0.046	0.070	0.063	0.106	0.046	0.086
500	5	0.062	0.054	0.052	0.144	0.067	0.131
500	10	0.056	0.066	0.067	0.168	0.055	0.135
1000	3	0.049	0.072	0.064	0.171	0.055	0.146
1000	5	0.053	0.068	0.075	0.260	0.065	0.209
1000	10	0.048	0.047	0.051	0.328	0.042	0.260

Notes: Monte Carlo experiments based on $R = 1,000$ replications. Bootstrap implementation uses $B = 200$ bootstrap replications.

Table 4: $\nu \sim skew - normal(shape = 10), \mu \sim N(0,1)$ and $\nu \sim N(0,1), \mu \sim skew - normal(shape = 10)$

$\nu \sim skew - normal(shape = 10), \mu \sim N(0,1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.780	0.069	0.673	0.051	0.063	0.050
100	5	0.999	0.193	0.999	0.055	0.073	0.062
100	10	1.000	0.506	1.000	0.048	0.060	0.046
200	3	0.967	0.178	0.960	0.058	0.060	0.058
200	5	1.000	0.422	1.000	0.043	0.066	0.067
200	10	1.000	0.885	1.000	0.042	0.090	0.077
500	3	1.000	0.527	1.000	0.059	0.059	0.066
500	5	1.000	0.908	1.000	0.058	0.079	0.078
500	10	1.000	0.997	1.000	0.049	0.078	0.069
1000	3	1.000	0.901	1.000	0.058	0.057	0.059
1000	5	1.000	0.999	1.000	0.064	0.066	0.068
1000	10	1.000	1.000	1.000	0.050	0.062	0.069
$\nu \sim N(0,1), \mu \sim skew - normal(shape = 10)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.052	0.092	0.080	0.596	0.036	0.464
100	5	0.051	0.073	0.065	0.761	0.044	0.616
100	10	0.041	0.079	0.063	0.855	0.049	0.733
200	3	0.054	0.072	0.073	0.917	0.065	0.836
200	5	0.051	0.082	0.070	0.971	0.057	0.914
200	10	0.058	0.056	0.060	0.993	0.078	0.968
500	3	0.046	0.068	0.062	1.000	0.154	0.999
500	5	0.061	0.053	0.051	0.999	0.182	0.998
500	10	0.058	0.067	0.066	1.000	0.217	1.000
1000	3	0.050	0.070	0.066	1.000	0.433	1.000
1000	5	0.054	0.068	0.074	1.000	0.458	1.000
1000	10	0.049	0.048	0.051	1.000	0.506	1.000

Notes: Monte Carlo experiments based on $R = 1,000$ replications. Bootstrap implementation uses $B = 200$ bootstrap replications.

Table 5: $\nu \sim skew - t_9(shape = 1), \mu \sim N(0,1)$ and $\nu \sim N(0,1), \mu \sim skew - t_9(shape = 1)$

$\nu \sim skew - t_9(shape = 1), \mu \sim N(0,1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.287	0.339	0.475	0.060	0.081	0.069
100	5	0.509	0.478	0.645	0.049	0.064	0.058
100	10	0.713	0.623	0.779	0.049	0.085	0.066
200	3	0.496	0.516	0.660	0.048	0.072	0.066
200	5	0.716	0.630	0.779	0.047	0.088	0.079
200	10	0.805	0.664	0.836	0.059	0.081	0.074
500	3	0.744	0.646	0.803	0.054	0.070	0.070
500	5	0.832	0.667	0.847	0.064	0.085	0.083
500	10	0.903	0.752	0.904	0.048	0.081	0.061
1000	3	0.828	0.686	0.844	0.061	0.039	0.058
1000	5	0.909	0.763	0.913	0.046	0.059	0.062
1000	10	0.926	0.785	0.924	0.047	0.073	0.063
$\nu \sim N(0,1), \mu \sim skew - t_9(shape = 1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.038	0.075	0.055	0.190	0.171	0.271
100	5	0.047	0.085	0.074	0.163	0.169	0.261
100	10	0.048	0.062	0.058	0.206	0.221	0.316
200	3	0.043	0.071	0.050	0.295	0.294	0.441
200	5	0.061	0.077	0.075	0.292	0.318	0.448
200	10	0.056	0.059	0.056	0.338	0.346	0.487
500	3	0.041	0.079	0.064	0.511	0.491	0.650
500	5	0.053	0.045	0.046	0.520	0.490	0.648
500	10	0.042	0.055	0.048	0.546	0.534	0.677
1000	3	0.052	0.067	0.072	0.695	0.591	0.785
1000	5	0.045	0.052	0.054	0.691	0.573	0.763
1000	10	0.062	0.065	0.066	0.686	0.605	0.774

Notes: Monte Carlo experiments based on $R = 1,000$ replications. Bootstrap implementation uses $B = 200$ bootstrap replications.

Table 6: $\nu \sim skew{-}t_9(shape = 10), \mu \sim N(0,1)$ and $\nu \sim N(0,1), \mu \sim skew{-}t_9(shape = 10)$

$\nu \sim skew - t_9(shape = 10), \mu \sim N(0,1)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.639	0.392	0.685	0.062	0.094	0.089
100	5	0.778	0.458	0.781	0.038	0.074	0.052
100	10	0.831	0.536	0.826	0.054	0.093	0.079
200	3	0.761	0.481	0.761	0.058	0.083	0.088
200	5	0.861	0.560	0.854	0.044	0.104	0.092
200	10	0.896	0.629	0.878	0.048	0.093	0.082
500	3	0.870	0.555	0.857	0.062	0.065	0.079
500	5	0.907	0.666	0.894	0.049	0.094	0.088
500	10	0.935	0.733	0.924	0.042	0.078	0.054
1000	3	0.900	0.668	0.888	0.063	0.058	0.072
1000	5	0.953	0.737	0.940	0.052	0.064	0.063
1000	10	0.948	0.754	0.941	0.046	0.077	0.069
$\nu \sim N(0,1), \mu \sim skew - t_9(shape = 10)$							
		Remainder component ν			Individual component μ		
N	T	Skewness	Kurtosis	Sk&Ku	Skewness	Kurtosis	Sk&Ku
100	3	0.047	0.075	0.064	0.629	0.041	0.471
100	5	0.063	0.084	0.078	0.750	0.052	0.630
100	10	0.069	0.076	0.077	0.835	0.050	0.734
200	3	0.050	0.097	0.091	0.922	0.072	0.838
200	5	0.051	0.065	0.069	0.973	0.062	0.939
200	10	0.047	0.082	0.066	0.987	0.056	0.971
500	3	0.049	0.079	0.068	0.999	0.183	0.996
500	5	0.069	0.055	0.058	1.000	0.177	0.999
500	10	0.062	0.062	0.070	1.000	0.200	1.000
1000	3	0.048	0.068	0.059	1.000	0.448	1.000
1000	5	0.035	0.053	0.049	1.000	0.439	1.000
1000	10	0.049	0.050	0.056	1.000	0.539	1.000

Notes: Monte Carlo experiments based on $R = 1,000$ replications. Bootstrap implementation uses $B = 200$ bootstrap replications.