Supplemental Appendices for

Normal but skew?

Dante Amengual

CEMFI

<amengual@cemfi.es>

Xinyue Bei

Duke University

<xinyue.bei@duke.edu>

Enrique Sentana

CEMFI

<sentana@cemfi.es>

May 2021

B Computational details of the simulations

We simulate n random draws from the multivariate skew normal distribution in (1) using the following rejection sampling method. First, we simulate $\mathbf{x} \sim N[\boldsymbol{\varphi}_M, \boldsymbol{\Sigma}(\boldsymbol{\varphi}_V)]$ together with an independent scalar random variable u with a uniform distribution between 0 and 1. If $u \leq \Phi\left[\boldsymbol{\vartheta}'\mathrm{dg}^{-1/2}(\boldsymbol{\varphi}_D)(\mathbf{x}-\boldsymbol{\varphi}_M)\right]$, then $\mathbf{y}=\mathbf{x}$, otherwise we discard it.

Arellano-Valle and Azzalini (2008) introduce an alternative parametrization of the multivariate skew normal distribution, which they call the central parametrization, such that the parameters of interest coincide with the means, variance and covariances of the observed variables, as well as their marginal skewness coefficients. They go from the original parametrization $(\boldsymbol{\varphi}_M, \boldsymbol{\varphi}_V, \boldsymbol{\vartheta})$ to the central one in two steps. First, they consider an intermediate vector of parameters such that

$$\begin{aligned}
\boldsymbol{\mu} &= E(\mathbf{y}) = \boldsymbol{\varphi}_M + \boldsymbol{\tau}, \\
\boldsymbol{\Upsilon}(\boldsymbol{\upsilon}) &= V(\mathbf{y}) = \boldsymbol{\Sigma}(\boldsymbol{\varphi}_V) - \boldsymbol{\tau}\boldsymbol{\tau}', \\
\boldsymbol{\tau} &= \sqrt{\frac{2}{\pi}}\mathrm{dg}^{1/2}(\boldsymbol{\varphi}_D)\boldsymbol{\delta}
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{\delta} &= [1+\boldsymbol{\vartheta}'\boldsymbol{\Psi}(\boldsymbol{\varphi}_V)\boldsymbol{\vartheta}]^{-1/2}\boldsymbol{\Psi}(\boldsymbol{\varphi}_V)\boldsymbol{\vartheta}, \\
\boldsymbol{\Psi}(\boldsymbol{\varphi}_V) &= \mathrm{dg}^{-1/2}(\boldsymbol{\varphi}_D)\boldsymbol{\Sigma}(\boldsymbol{\varphi}_V)\mathrm{dg}^{-1/2}(\boldsymbol{\varphi}_D)
\end{aligned}$$

and $\boldsymbol{\upsilon} = (\boldsymbol{\upsilon}_D', \boldsymbol{\upsilon}_L')'$, with $\boldsymbol{\upsilon}_D = vecd[\boldsymbol{\Upsilon}(\boldsymbol{\upsilon})]$ and $\boldsymbol{\upsilon}_L = vecl[\boldsymbol{\Upsilon}(\boldsymbol{\upsilon})]$. This reparametrization is a one-to-one mapping with a non-zero Jacobian determinant even at the Gaussian null. In addition, it is easy to prove that the scores corresponding to $\boldsymbol{\mu}$ evaluated at $\boldsymbol{\tau}=\mathbf{0}$ coincide with the scores corresponding to $\boldsymbol{\varphi}_M$ evaluated at $\boldsymbol{\vartheta}=\mathbf{0}$, the same being true of the scores for $\boldsymbol{\upsilon}$ and $\boldsymbol{\varphi}_V$. This is not entirely surprising in view of the fact that $\boldsymbol{\varphi}_M$ and $\boldsymbol{\Sigma}(\boldsymbol{\varphi}_V)$ directly yield $E(\mathbf{y})$ and $V(\mathbf{y})$ under normality. In contrast, all the elements of the score vector and Hessian matrix corresponding to $\boldsymbol{\tau}$ are 0 when evaluated at $\boldsymbol{\tau}=\mathbf{0}$, thereby achieving the goal of confining the singularities to those elements, as in the proof of Proposition 1. Nevertheless, the third derivatives are no longer 0. Specifically,

$$\left.\frac{\partial^3 l}{\partial \tau_k^3}\right|_{\boldsymbol{\tau}=\mathbf{0}} = \frac{4-\pi}{2}H_{kkk}[\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Upsilon}(\boldsymbol{\upsilon})] + \frac{12}{(1-R_k^2)\upsilon_{D,k}}\left. s_{\mu_k}\right|_{\boldsymbol{\tau}=\mathbf{0}}, \tag{B1}$$

where $H_{kkk}[\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Upsilon}(\boldsymbol{\upsilon})]$ is one of the $K(K+1)(K+2)/6$ third-order multivariate Hermite polynomials in (3) and R_k^2 is the coefficient of determination in the regression of y_k on a constant and the remaining elements of $\mathbf{y}$.

Next, Arellano-Valle and Azzalini (2008) replace each τ_k with the corresponding marginal skewness coefficient

$$\gamma_k = \frac{E(y_k-\mu_k)^3}{\upsilon_{D,k}^{3/2}} = \frac{4-\pi}{2}\left(\frac{\tau_k}{\sqrt{\upsilon_{D,k}}}\right)^3.$$

The problem with this reparametrization is that its first and second derivatives are 0 under the Gaussian null, but this is precisely the trick that Lee and Chesher (1986) used to re-interpret their extremum test as an LM test in the case of a single parameter. Specifically, after applying L'Hopital's rule twice, the score of γ_k evaluated at $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_K)'=\mathbf{0}$ is

$$\left.\frac{\partial l}{\partial \gamma_i}\right|_{\boldsymbol{\gamma}=\mathbf{0}} = \frac{\upsilon_{D,k}}{6}H_{kkk}[\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Upsilon}(\boldsymbol{\upsilon})] + \frac{4\upsilon_{D,k}}{(4-\pi)(1-R_k^2)}\left. s_{\mu_k}\right|_{\boldsymbol{\gamma}=\mathbf{0}}, \tag{B2}$$

which is proportional to (B1). Once we purge these derivatives from the effects of estimating the sample mean vector and covariance matrix by regressing them on the scores with respect to $\boldsymbol{\mu}$ and $\boldsymbol{\upsilon}$ and retaining the residuals, we end up with the moment test based on $H_{kkk}[\mathbf{y};\boldsymbol{\mu},\boldsymbol{\Upsilon}(\boldsymbol{\upsilon})]$ for $k=1,\ldots,K$.

Clearly, this procedure ignores all the other $K(K-1)(K+4)/6$ third cross-derivatives of $\boldsymbol{\tau}$ and $\boldsymbol{\gamma}$, which depend on the remaining third-order multivariate Hermite polynomials in (3).

Figure A1: Alignment of GET and LR under the Gaussian null

Figure A1a: Bivariate case

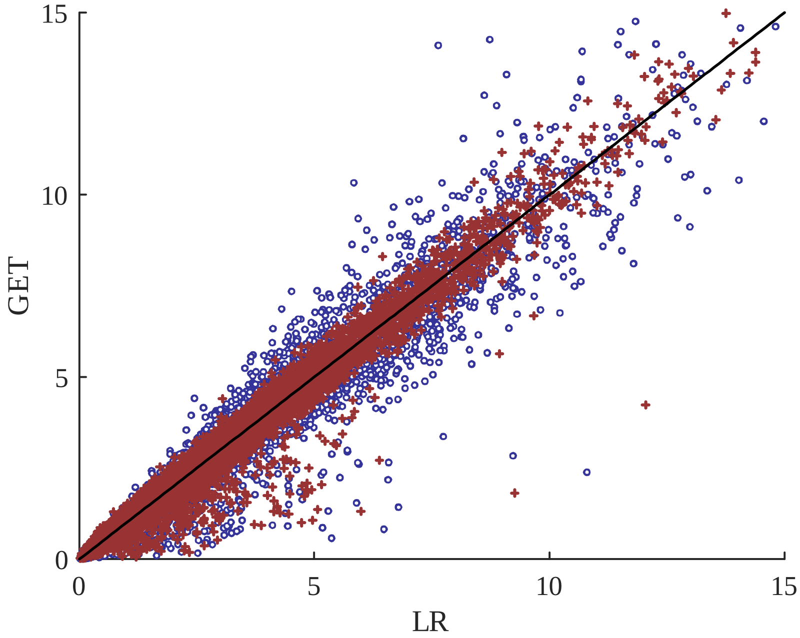
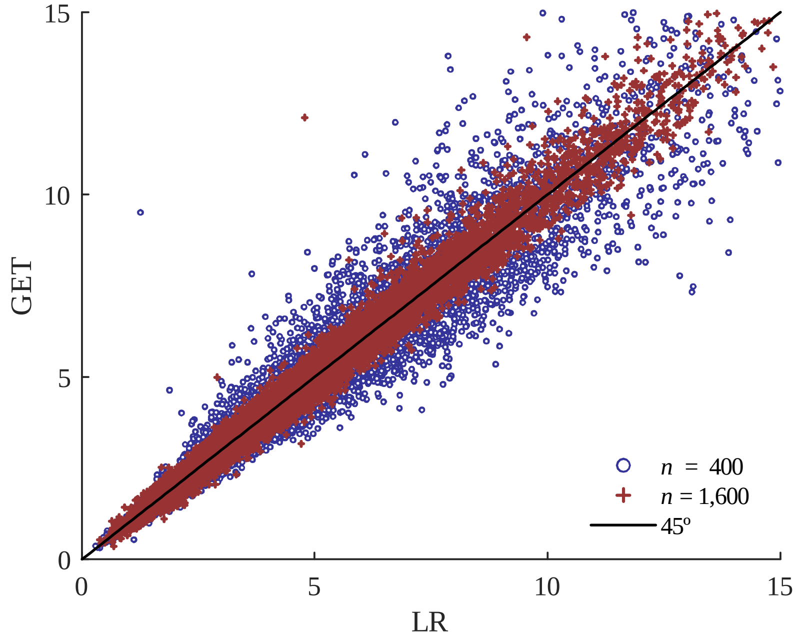


Figure A1b: Trivariate case



Notes: Scatter plots of the GET and LR test statistics based on 10,000 samples. Upper and lower panels display results for bivariate and trivariate models, respectively. The true mean and covariance matrix of the simulated Gaussian data are set to $\mathbf{0}$ and $\mathbf{I}_k$, while the mean and variance parameters $\boldsymbol{\varphi}_M$ and $\boldsymbol{\varphi}_V$ are estimated under the null using the sample mean and covariance matrix, respectively.