Supplementary material for: ‘Hotel pricing, stochastic demand and Covid-19’

*
Andrea Guizzardi
Department of Statistical Sciences,
University of Bologna,
Via delle Belle Arti 41, 40126 Bologna, Italy,
e-mail: andrea.guizzardi@unibo.it

Luca Vincenzo Ballestra
Department of Statistical Sciences,
University of Bologna,
Via delle Belle Arti 41, 40126 Bologna, Italy,
e-mail: luca.ballestra@unibo.it

and

Enzo D’Innocenzo*
Department of Econometrics and Data Science,
Vrije Universiteit Amsterdam
De Boelelaan 1105, 1081 HV Amsterdam, Netherlands
e-mail: e.dinnocenzo@vu.nl

October 21, 2022

*Corresponding author.

1 Specification and estimation of the models

1.1 The econometric models

Let $p(\epsilon_{t,k}|\mathcal{F}^k_{t-1}, \boldsymbol{\vartheta}_k)$ denote the conditional probability density function of the price shock, where $\mathcal{F}^k_{t-1} = \sigma\{\epsilon_{t-1,k}, \epsilon_{t-2,k}, \dots\}$ is the filtration generated by $\{\epsilon_{t,k}\}$ and $\boldsymbol{\vartheta}_k$ is the vector of unknown parameters of the conditional density. Depending on which of the three models we consider, $\boldsymbol{\vartheta}_k$ will vary accordingly.

As a first step, we assume that $\epsilon_{t,k}$ is normally distributed with location (mean) μ_k and scale (standard deviation) $\varphi_k \in \mathbb{R}^+$, and thus $p(\epsilon_{t,k}|\mathcal{F}^k_{t-1}, \boldsymbol{\vartheta}_k) = \mathcal{N}(\mu_k, \varphi_k^2)$ and $\boldsymbol{\vartheta}_k = (\mu_k, \varphi_k)^\top$.

Furthermore, following Fernandez and Steel (1998) and Azzalini (2013), we also model $\epsilon_{t,k}$ as a static skew-t distribution with $\nu_k \in (1, +\infty)$ degrees of freedom and skewness parameter $\gamma_k \in \mathbb{R}^+$, i.e., $p(\epsilon_{t,k}|\mathcal{F}^k_{t-1}, \boldsymbol{\vartheta}_k) = \textit{Skew}\text{-t}(\mu_k, \varphi_k^2, \nu_k, \gamma_k)$, and $\boldsymbol{\vartheta}_k = (\mu_k, \varphi_k^2, \nu_k, \gamma_k)^\top$. Note that the constraint $\nu_k > 1$ is imposed to ensure that $\mathbb{E}[\epsilon_{t,k}] < \infty$.

The resulting models are labeled as:

$$\textbf{\textit{Model 1}}: \qquad P_{t,0} = b_k P_{t,k} + \epsilon_{t,k}, \qquad \epsilon_{t,k} \overset{IID}{\sim} \mathcal{N}(\mu_k, \varphi_k^2), \tag{1}$$

$$\textbf{\textit{Model 2}}: \qquad P_{t,0} = b_k P_{t,k} + \epsilon_{t,k}, \qquad \epsilon_{t,k} \overset{IID}{\sim} \textit{Skew}\text{-t}(\mu_k, \varphi_k^2, \nu_k, \gamma_k). \tag{2}$$

If we assume that the four parameters of the probability distribution of $\epsilon_{t,k}$, i.e., location, scale, kurtosis and skewness, are time-varying (i.e., they vary according to seasonality), we can model them according to the score-driven approach proposed by Creal et al. (2011) and Harvey (2013). Then, we suggest the following model:

$$\textbf{\textit{Model 3}}: \qquad P_{t,0} = b_k P_{t,k} + \epsilon_{t,k}, \qquad \epsilon_{t,k}|\mathcal{F}^k_{t-1} \sim \textit{Skew}\text{-t}\,(\mu_{t,k}, \varphi_{t,k}^2, \nu_{t,k}, \gamma_{t,k}). \tag{3}$$

It is important to observe that $\varphi_{t,k}$, $\gamma_{t,k}$ and $\nu_{t,k}$ must lay in proper spaces for every t. To guarantee this, we transform them, via suitable link functions, to unconstrained parameters that are allowed to float freely. Specifically, to ensure that $\varphi_{t,k}$ and $\gamma_{t,k} \in \mathbb{R}_+$, we use the exponential link functions $\varphi_{t,k} = \exp\{\lambda_{t,k}\}$ and $\gamma_{t,k} = \exp\{\xi_{t,k}\}$, where $\lambda_{t,k}$, $\xi_{t,k} \in \mathbb{R}$. Finally, for the degrees of freedom $\nu_{t,k}$ we opt for the transformation $\nu_{t,k} = 1 + \exp\{\psi_{t,k}\}$, where $\psi_{t,k} \in \mathbb{R}$.

Therefore, we consider the change of variables

$$\boldsymbol{\Lambda}(\boldsymbol{f}_{t,k}) = \begin{bmatrix} \mu_{t,k} \\ \exp\{\lambda_{t,k}\} \\ 1 + \exp\{\psi_{t,k}\} \\ \exp\{\xi_{t,k}\} \end{bmatrix}, \tag{4}$$

where $\boldsymbol{f}_{t,k} = (\mu_{t,k}, \lambda_{t,k}, \psi_{t,k}, \xi_{t,k})^\top$. Then, according to Creal et al. (2011) and Harvey (2013), we update the distribution parameters by using the following first-order vector recursion

$$\boldsymbol{f}_{t+1,k} = \boldsymbol{\delta}_k + \boldsymbol{\Phi}_k \boldsymbol{f}_{t,k} + \boldsymbol{K}_k \boldsymbol{s}_{t,k}, \tag{5}$$

where $\boldsymbol{s}_{t,k}$ is the unconstrained conditional scaled score, $\boldsymbol{\delta}_k = (\delta_{\mu_k}, \delta_{\lambda_k}, \delta_{\psi_k}, \delta_{\xi_k})^\top \in \mathbb{R}^4$ is a vector of intercepts and $\boldsymbol{\Phi}_k \in \mathbb{R}^{4\times4}$ and $\boldsymbol{K}_k \in \mathbb{R}^{4\times4}$ are diagonal matrices with $\text{diag}(\boldsymbol{\Phi}_k) = (\phi_{\mu_k}, \phi_{\lambda_k}, \phi_{\psi_k}, \phi_{\xi_k})^\top$ and $\text{diag}(\boldsymbol{K}_k) = (\kappa_{\mu_k}, \kappa_{\lambda_k}, \kappa_{\psi_k}, \kappa_{\xi_k})^\top$. Therefore, the vector of parameter for *Model 3* is $\boldsymbol{\vartheta}_k =$

$(\delta_{\mu_k}, \delta_{\lambda_k}, \delta_{\psi_k}, \delta_{\xi_k}, \phi_{\mu_k}, \phi_{\lambda_k}, \phi_{\psi_k}, \phi_{\xi_k}, \kappa_{\mu_k}, \kappa_{\lambda_k}, \kappa_{\psi_k}, \kappa_{\xi_k})^\top$. As a standard approach, we impose the constraints $|\phi_{\mu_k}| < 1$, $|\phi_{\lambda_k}| < 1$, $|\phi_{\psi_k}| < 1$ and $|\phi_{\xi_k}| < 1$ to keep the recursion (5) stable.

To calculate the unrestricted score in (5), we need to consider the conditional log-density of $\epsilon_{t,k}$, which we parametrize as in Harvey and Sucarrat (2014):

$$\begin{aligned}\ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k) = \ln 2 - \ln(\gamma_{t,k} + 1/\gamma_{t,k}) + \ln\Gamma\Big(\frac{\nu_{t,k}+1}{2}\Big) - \ln\Gamma\Big(\frac{\nu_{t,k}}{2}\Big) - \frac{1}{2}\ln\pi \\ - \frac{1}{2}\ln\varphi_{t,k}^2 - \frac{\nu_{t,k}+1}{2}\ln\left(1 + \frac{(\epsilon_{t,k}-\mu_{t,k})^2}{\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}\nu_{t,k}\varphi_{t,k}^2}\right).\end{aligned} \tag{6}$$

The driving-force $\boldsymbol{s}_{t,k}$ is computed as follows

$$\boldsymbol{s}_{t,k} = \boldsymbol{J}_{t,k}^\top \boldsymbol{\nabla}_{t,k}, \tag{7}$$

where $\boldsymbol{\nabla}_{t,k} = \frac{\partial \ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k)}{\partial \boldsymbol{f}_{t,k}}$ is the score vector of the predictive log-density in (6), and $\boldsymbol{J}_{t,k} = \frac{\partial \boldsymbol{\Lambda}(\boldsymbol{f}_{t,k})}{\partial \boldsymbol{f}_{t,k}^\top}$ is the Jacobian matrix of the nonlinear mapping $\boldsymbol{\Lambda}(\,\cdot\,)$.

By taking derivatives in (4), we have

$$\boldsymbol{J}_{t,k} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \exp\{\lambda_{t,k}\} & 0 & 0 \\ 0 & 0 & \exp\{\psi_{t,k}\} & 0 \\ 0 & 0 & 0 & \exp\{\xi_{t,k}\} \end{bmatrix},$$

whereas $\boldsymbol{\nabla}_{t,k} = (\nabla_{t,k}^\mu, \nabla_{t,k}^\varphi, \nabla_{t,k}^\nu, \nabla_{t,k}^\gamma)^\top$ in equation (7) is computed as follows

$$\nabla_{t,k}^\mu = \frac{\partial \ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k)}{\partial \mu_{t,k}} = \begin{cases} \frac{(\nu_{t,k}+1)(\epsilon_{t,k}-\mu_{t,k})/(\gamma_{t,k}^{-2}\nu_{t,k}\varphi_{t,k}^2)}{1+(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{-2}\nu_{t,k}\varphi_{t,k}^2)} & \text{if } (\epsilon_{t,k}-\mu_{t,k}) \in (-\infty, 0), \\ \frac{(\nu_{t,k}+1)(\epsilon_{t,k}-\mu_{t,k})/(\gamma_{t,k}^{2}\nu_{t,k}\varphi_{t,k}^2)}{1+(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2}\nu_{t,k}\varphi_{t,k}^2)} & \text{if } (\epsilon_{t,k}-\mu_{t,k}) \in [0, +\infty), \end{cases}$$

$$\nabla_{t,k}^\varphi = \frac{\partial \ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k)}{\partial \varphi_{t,k}} = \begin{cases} \frac{1}{\varphi_{t,k}}\left(\frac{(\nu_{t,k}+1)(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{-2}\nu_{t,k}\varphi_{t,k}^2)}{1+(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{-2}\nu_{t,k}\varphi_{t,k}^2)} - 1\right) & \text{if } (\epsilon_{t,k}-\mu_{t,k}) \in (-\infty, 0), \\ \frac{1}{\varphi_{t,k}}\left(\frac{(\nu_{t,k}+1)(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2}\nu_{t,k}\varphi_{t,k}^2)}{1+(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2}\nu_{t,k}\varphi_{t,k}^2)} - 1\right) & \text{if } (\epsilon_{t,k}-\mu_{t,k}) \in [0, +\infty), \end{cases}$$

$$\begin{aligned}\nabla_{t,k}^\nu = \frac{\partial \ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k)}{\partial \nu_{t,k}} = \frac{1}{2}\Bigg[\Psi\left(\frac{\nu_{t,k}+1}{2}\right) - \Psi\left(\frac{\nu_{t,k}}{2}\right) - \frac{1}{\nu_{t,k}} - \ln\left(1 + \frac{(\epsilon_{t,k}-\mu_{t,k})^2}{\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}\nu_{t,k}\varphi_{t,k}^2}\right) \\ + \frac{\nu_{t,k}+1}{\nu_{t,k}}\left(\frac{(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}\nu_{t,k}\varphi_{t,k}^2)}{1+(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}\nu_{t,k}\varphi_{t,k}^2)}\right)\Bigg],\end{aligned}$$

$$\nabla_{t,k}^\gamma = \frac{\partial \ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k)}{\partial \gamma_{t,k}} = \frac{1-\gamma_{t,k}^2}{\gamma_{t,k}^3+\gamma_{t,k}} + \mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})-1}$$

$$\times \frac{\nu_{t,k}+1}{\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}}\left(\frac{(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}\nu_{t,k}\varphi_{t,k}^2)}{1+(\epsilon_{t,k}-\mu_{t,k})^2/(\gamma_{t,k}^{2\mathrm{sgn}(\epsilon_{t,k}-\mu_{t,k})}\nu_{t,k}\varphi_{t,k}^2)}\right),$$

where $\Psi(x) = \frac{d}{dx}\ln\Gamma(x)$ is the digamma function, see Abramowitz and Stegun (1964).

1.2 Model estimation

All the models described in the previous Subsections can be estimated by maximum likelihood (ML). For each of them, in addition to the vector of parameters $\boldsymbol{\vartheta}_k$, we also have to estimate the parameter b_k. The regressor $P_{t,k}$ in (1), (2) and (3) is potentially endogenous. Endogeneity issues might occur due to missing variables, since prices can also vary with the room quality, which is not fully observable based on the information gathered from the Internet. Additionally, the last-minute price adjustment $\epsilon_{t,k}$ is also affected by the unforeseen demand on both the OTA and the other distribution channels (which determines the offer adjustment $\varsigma_{t,1}$), which, in turn, might depend on the fixed early booking price $P_{t,k}$.

To handle possible endogeneity, we adopt the nonlinear instrumental variable (IV) approach as in Hansen et al. (2010), with a two-step estimation procedure. First, we obtain a preliminary estimation of $\tilde{b}_k$. For *Model 1* and *Model 2* we use a standard two-stage least squares estimator where, for a given hotel i, the price $P_{t,k}$ is instrumented by the average of the prices $P_{t,k}^{(\cdot)}$ posted by the other hotels:

$$z_{t,k} = \frac{1}{L-1}\sum_{\substack{l=1\\ l\neq i}}^{L} P_{t,k}^{(l)}, \tag{8}$$

where L is the number of the hotels in the price quartile to which hotel i belongs. In particular we define quartiles by ordering hotels according to their median price (all t and k). Instead, for *Model 3*, $\tilde{b}_k$ is computed by standard ML.

Then, we calculate the residuals $\tilde{\epsilon}_{t,k} = P_{t,0} - \tilde{b}_k P_{t,k}$, and a quasi-ML estimation of $\boldsymbol{\vartheta}_k$ is obtained as follows:

$$\tilde{\boldsymbol{\vartheta}}_k = \arg\max_{\boldsymbol{\vartheta}_k} \sum_{t=1}^{T} \ln p(\tilde{\epsilon}_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k).$$

Finally, following Hansen et al. (2010), the nonlinear IV estimator of b_k is given by

$$\hat{b}_k = \arg\min_{b_k} \frac{[\hat{g}(b_k)]^2}{\hat{\mathcal{Q}}_k},$$

where $\hat{\mathcal{Q}}_k = \sum_{t=1}^{T} z_{t,k}^2$, $z_{t,k}$ is the instrumental variable defined in (8) and

$$\hat{g}(b_k) = \sum_{t=1}^{T} z_{t,k}\rho(P_{t,0} - b_k P_{t,k}, \tilde{\boldsymbol{\vartheta}}_k), \qquad \rho(\epsilon_{t,k}, \boldsymbol{\vartheta}_k) = \frac{\partial \ln p(\epsilon_{t,k}|\mathcal{F}_{t-1}^k, \boldsymbol{\vartheta}_k)}{\partial \epsilon_{t,k}}.$$

We test the significance of the estimated parameters using the robust standard errors derived in Hansen et al. (2010).

2 Last-minute price distribution for six selected hotels

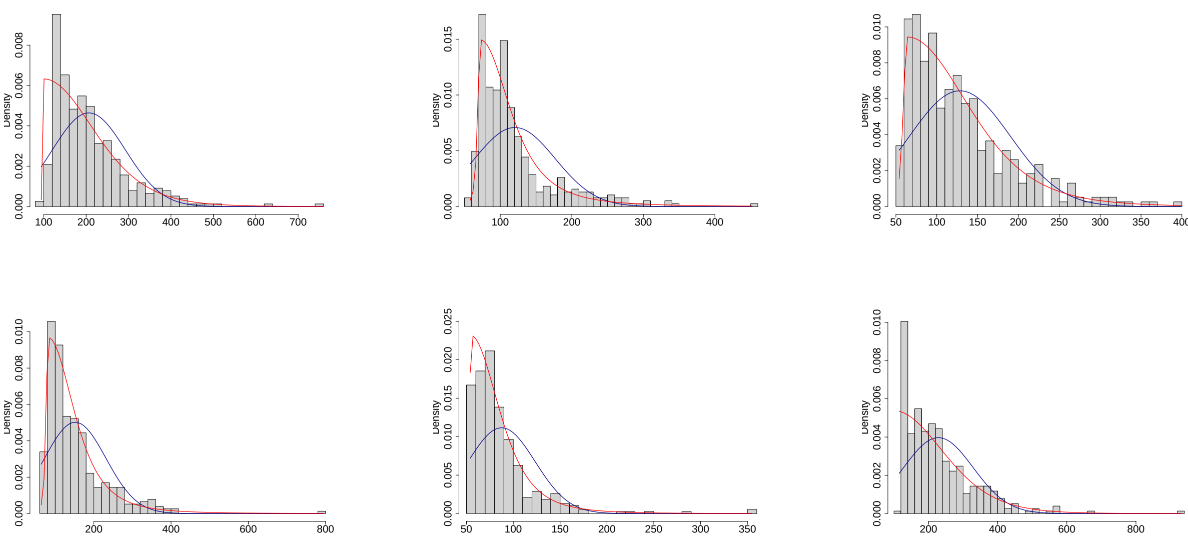

Figure 1: Price histogram, pre-Covid-19, skew-t kernel (red line), and Gaussian kernel (blue line).

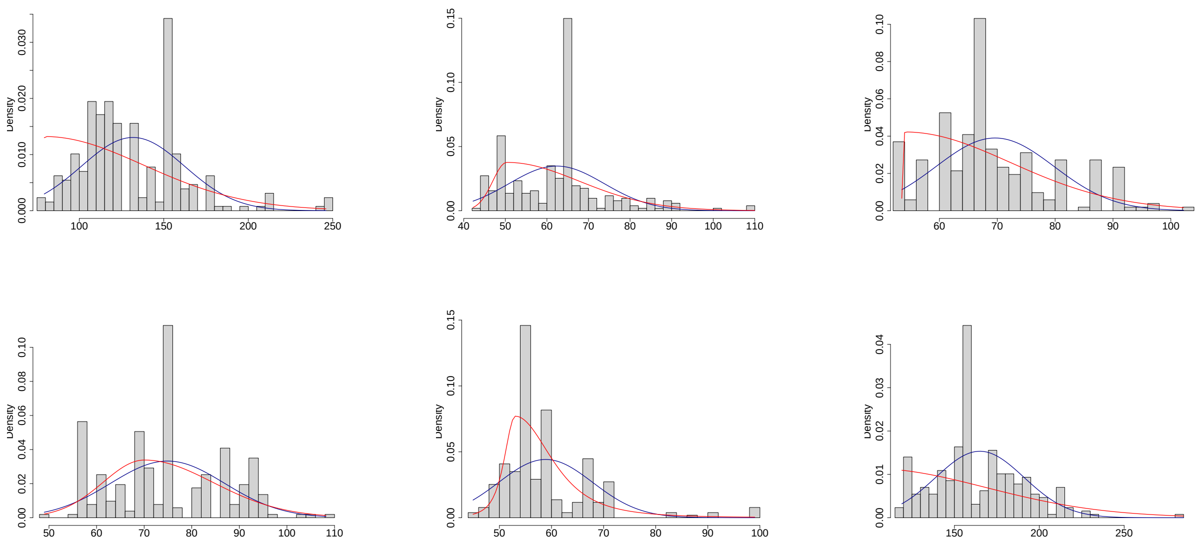

Figure 2: Price histogram, during-Covid-19, skew-t kernel (red line), and Gaussian kernel (blue line).

References

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York.

Azzalini, A. (2013). *The Skew-Normal and Related Families.* Institute of Mathematical Statistics Monographs. Cambridge University Press.

Creal, D., Koopman, S. J., and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29(4):552–563.

Fernandez, C. and Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371.

Hansen, C., McDonald, J. B., and Newey, W. K. (2010). Instrumental variables estimation with flexible distributions. *Journal of Business & Economic Statistics*, 28(1):13–25.

Harvey, A. and Sucarrat, G. (2014). EGARCH models with fat tails, skewness and leverage. *Computational Statistics* & *Data Analysis*, 76:320 – 338.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails.* Econometric Society Monograph - Cambridge University Press.