CENTRE FOR APPLIED MACROECONOMICS AND COMMODITY PRICES (CAMP)

CAMP Working Paper Series
No 5/2020

Large Time-Varying Volatility Models for Electricity Prices

Angelica Gianfreda, Francesco Ravazzolo and Luca Rossini

Large Time-Varying Volatility Models for Electricity Prices*

Angelica Gianfreda[a,d] Francesco Ravazzolo[b,e] Luca Rossini[c,f]

[a]University of Modena and Reggio Emilia, Modena, Italy [b]Free University of Bozen-Bolzano, Italy

[c]Vrije Universiteit Amsterdam, The Netherlands [d]EMG, London Business School, UK

[e]CAMP, BI Norwegian Business School, Norway [f]Ca' Foscari University of Venice, Italy

July 2, 2020

Abstract

We study the importance of time-varying volatility in modelling hourly electricity prices when fundamental drivers are included in the estimation. This allows us to contribute to the literature of large Bayesian VARs by using well-known time series models in a huge dimension for the matrix of coefficients. Based on novel Bayesian techniques, we exploit the importance of both Gaussian and non-Gaussian error terms in stochastic volatility. We find that by using regressors as fuels prices, forecasted demand and forecasted renewable energy is essential in order to properly capture the volatility of these prices. Moreover, we show that the time-varying volatility models outperform the constant volatility models in both the in-sample model-fit and the out-of-sample forecasting performance.

Keywords: Electricity; Hourly Prices; Renewable Energy Sources; Non-Gaussian; Stochastic-Volatility; Forecasting.

*The authors gratefully acknowledge Matteo Iacopini and Jamie Cross for their useful feedback. This paper is part of the research activities at the Free University of Bozen-Bolzano funded by Europe Energy S.p.A. This paper is part of the research activities at the Centre for Applied Macroeconomics and Commodity Prices (CAMP) at the BI Norwegian Business School. This research used the SCSCF multiprocessor cluster system at Ca' Foscari University of Venice. Luca Rossini acknowledges the financial support from the EU Horizon 2020 programme under the Marie Skłodowska-Curie scheme (grant agreement no.796902).

1 Introduction

Due to the recent advent and availability of big data sets for forecasting important macroeconomic variables, scholars and researchers have focused on the macroeconometrics of time series analysis, see Huber and Feldkircher (2019); Koop et al. (2019); Huber et al. (2020); Chan (2020a) among others. However, it is only recently that we have observed a growing interest in the multivariate analysis of electricity prices and the influence that variable and unpredictable renewable energy sources (such as wind and, to lesser extent, solar photovoltaic energy) can exercise on the determination of these prices. Indeed, the vast majority of the literature on electricity prices focuses on univariate models for individual hours of the day and only a few papers have recently considered multivariate specifications (Raviv et al., 2015; Gianfreda et al., 2020). To the best of our knowledge, the analysis of multivariate models with time-varying volatility has not been deeply investigated and we aim at filling also this gap adopting Bayesian approaches.

Recently, the question of time instability has been posed to account, for instance, for large drops in demand that can suddenly occur (as during the current pandemic) or for negative wholesale electricity prices. Then, it has become clear that electricity, as well as interest rates and oil prices, can become negative and attract the attention of all economists to understand the economic consequences; as for those implied by the dynamics of WTI prices in April 2020.

In the energy markets, the phenomenon of negative prices - when allowed to occur - has become more frequent due to the increasing share of electricity generated from renewable energy sources (RES). Indeed, worldwide energy policies have supported, and they are still fostering, green generation to reduce carbon emissions and mitigate the climate change. From a technical point of view, RES have induced prices to be null or even negative. Indeed, being a clean energy source with null marginal cost of production, RES enters the supply curve before the other thermal conventional technologies with higher marginal costs (this is known as the *merit order*). Consequently, the supply curve is shifted towards the right by the amount of generation produced by intermittent renewable energy sources. This reduces the equilibrium price that, in some markets like Germany, can become negative when extremely high RES generation is coupled by low levels of demand.

However, renewables (as wind, hydro and solar to less extent) are variable, intermittent and not easily predictable since they are strictly related to weather conditions. If the wind blows and/or the sun shines, green and economic generation satisfies the demand and electricity prices are low (or negative); otherwise, demanded electricity is covered with more expensive thermal conventional plants running with fossil fuels. Consequently, there are recent and increasing concerns about global warming, which will inevitably amplify these patterns and it will make demand as well as supply varying over time with greater uncertainty; see for instance Damm et al. (2017).

Econometrically, all these issues support the application of time-varying models, as the multivariate models with time-varying volatility. Dimension is further expanded when we consider the needs to include past information to account for dependencies and strong seasonality affecting electricity markets. Indeed, demand is heavily influenced by industrial activities with different dynamics over the week and the year; and prices are strictly related among themselves and their past observations.

The most challenging point for the use of multivariate time series models is the huge amount of information available in the exogenous variables and in their lagged values. In fact, in order to estimate the simplest model, we need a lagged representation for each hourly price depending on its past values observed on one, two and seven days before, plus the inclusion of dummy variables; hence, leading to 72 parameters for the 24 hourly prices with their previous 3 lags, plus 13 dummies variables for 12 months and the one indicating weekends and holidays. However, including for each hour also the exogenous variables for supply, that is RES and fossil fuels, and demand, the matrix of coefficients becomes of size 160 times 24 hours.

One can state that traditionally factor models are successfully used to handle large datasets, however the recent literature in large Bayesian Vector Autoregressive models (VAR) has provided valid and important alternatives. In the macroeconomic literature (Clark and Ravazzolo, 2015; Carriero et al., 2016, 2019; Chan, 2020a), the use of time-varying volatility is known to improve the full sample analysis by capturing the peak and booms and also to improve the forecasting accuracy. See Bauwens et al. (2013) for early research in time-varying volatility for electricity markets. Thus, following this stream of the literature, we have decided to expand the models proposed by Gianfreda et al. (2020) by adding stochastic volatility. Indeed in that paper, authors compare several univariate and multivariate frequentist and Bayesian models augmented with fundamental variables to predict hourly day-ahead electricity prices, but without considering the issue of volatility. In the energy literature, there are mixed results related to volatility, with the main ones arguing that time-varying volatility can be captured by the intermittent behaviour of renewable energy sources. As suggested by Karakatsani and Bunn (2010), the inclusion of fundamental drivers can eliminate volatility effects, being a sort of surrogate for omitted factors.

However, it must be emphasised that all previous research has considered only the univariate dimension and multivariate models have not been explored. The large dimensionality of time-varying volatility models applied to (electricity) hourly data[1] can result in sizeable estimation errors. To mitigate this, we follow recent macroeconomic literature and apply Bayesian estimation techniques.

Furthermore, to our knowledge, this paper is the first to provide a full sample multivariate analysis of the volatility of electricity prices, and the results refer to two European countries (Germany and Italy) that have a strong RES influx and, in the German case, allowing the formation of negative prices.[2] Besides electricity markets, these models will become of extreme interest in the future applications and consequent developments, since we have already assisted to the formation of negative interest rates in response to financial crises. Thus, the economic consequences of the global pandemic situation will enhance more the determination of negative (or extremely low) prices also for other commodities, as it has already occurred for crude oil.

From the computational and operational point of view, Carriero et al. (2019) and Chan (2020a) have recently developed Bayesian estimation methods for VARs with stochastic volatility that allow to reduce the computational timing, when the number of coefficients increase roughly. Therefore, following the specification in Chan and Eisenstat (2018) and Cross et al. (2020), we use a Bayesian VAR with stochastic volatility for showing the movements of the volatility in the electricity prices. Differently from Chan (2020a), we do not assume a stochastic volatility component constant across the variables, but we assume that the time-varying volatility changes across hours. Moreover, differently from Carriero et al. (2019) and Cross et al. (2020), we do not have any ordering problems in the estimation, because our variables follow a time dependence structure of consecutive hours in a day.

The main finding in our paper refers to the importance of modelling multivariate time-varying volatility in the variation of the electricity prices and, in particular, the fact that the assumption of constant volatility on average overestimates the time-varying volatility over time, thus leading to imprecise estimation. Moreover, the inclusion of fat tail error term in the stochastic volatility produces further improvements across central hours, where the estimation of different degrees of freedom is important.

In the paper, we provide also a forecasting exercise of the German and Italian electricity prices

[1]We remember that our models requires to estimate a (24×24) time-varying covariance matrix.
[2]We count 595 negative prices in our sample, corresponding to almost 2% of all observations.

from 2018 to 2019. We found evidence of strong improvements once we include forecasted demand and renewable energy sources in the analysis with respect to the baseline model with only a lag representation of the response variable. Furthermore, if we include time-varying volatility in the form of Gaussian or Student-t form, we observe strong improvements in Germany and moderate improvements in Italy. These improvements are available both in point and density forecasting, and also when focusing on quantile density forecasting for tail comparison. Therefore, as argued by Gianfreda and Bunn (2018), price asymmetries induced by wind generation are significant and attention shall be paid to their modelling. The presence of fat tail in the stochastic volatility is particularly emphasised in Germany during the central hours and across the different metrics considered. This phenomenon is particularly emphasised in this market since prices are free to fluctuate from a floor price of -500 €/MWh to a cap price of 3000€/MWh, whereas in other countries the floor is generally set to zero.

The paper proceeds as follows. We present the model structure used in the empirical study in Section 2 together with some details on the estimation methodology. Section 3 describes the data used. In Section 4, we consider the application of time-varying volatility models to hourly electricity prices. We first present the results based on a full sample estimation, then those based on different sample sizes. It also contains a recursive out-of-sample forecasting exercise to assess the performance of the Bayesian VAR with Stochastic volatility models by means of both point and density metrics. Finally, Section 5 concludes and briefly discusses some future research directions.

2 VARs with Stochastic Volatility

In this section, we outline the class of models we wish to compare. Firstly, we consider the most general model with time-varying volatility (Chan and Eisenstat, 2018; Cross et al., 2020) and then other models are specified as restricted versions of the general one. In particular, we highlight the differences between a Gaussian stochastic volatility model and a fat-tail stochastic volatility model, thus with the Student-t error term. Therefore, our multivariate specifications allow us to compare whether features such as intermittent and unpredictable supply and variable demand (and consequent negative prices) only increase volatility or actually change the (tail) distribution, as suggested by Gianfreda and Bunn (2018) in their analysis of univariate time series.

Let $\mathbf{y}_t = (y_{1t}, \ldots, y_{Ht})'$ denote the $(H \times 1)$ vector of day-ahead hourly electricity prices, with

$H = 24$. Consider the following vector autoregressive (VAR) model with stochastic-volatility (SV):

$$A_0 \mathbf{y}_t = B_1 \mathbf{y}_{t-1} + \ldots + B_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \Sigma_t), \tag{1}$$

where $B_1, \ldots, B_p$ are the $(H \times H)$ VAR matrix of coefficients; A_0 is a $(H \times H)$ lower triangular matrix with ones on the diagonal and Σ_t is a time-varying diagonal matrix of the form $\Sigma_t = \mathrm{diag}(\exp(h_{1t}), \ldots, \exp(h_{Ht}))$. Following Chan and Eisenstat (2018), we reformulate the model as follows:

$$\mathbf{y}_t = \tilde{X}_t \boldsymbol{\beta} + W_t \boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \Sigma_t), \tag{2}$$

where $\tilde{X}_t = I_H \otimes (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p})$ and W_t contains the appropriate elements of $\mathbf{y}_t$. Regarding the coefficients in $B_1, \ldots, B_p$ and A_0, we can split them in two different groups. The first group consists of $\boldsymbol{\beta}$, which is a $(k_\beta \times 1)$ vector containing the coefficients associated with the lagged observations, the dummy variables and the exogenous variables. On the other hand, the second group contains a $(k_\gamma \times 1)$ vector, $\boldsymbol{\gamma}$, of coefficients that characterizes the contemporaneous relations between the variables and it consists of the free elements of A_0 stacked by rows.

In particular, $k_\gamma = H(H-1)$, while the size of the vector of coefficients varies along the model specification. If we include in the specifications the lagged observations, then $k_\beta = H^2 p$. On the other hand, if we add a vector of dummies denoted by $\mathbf{d}_t = (d_{1t}, \ldots, d_{Kt})'$, where $(d_{1t}, \ldots, d_{12t})$ representing the twelve months of the year and d_{13t} representing Saturdays, Sundays and holidays for each country, hence $K = 13$; then $k_\beta = (Hp + K)H$.

The model in Eq. (2) can be written in a stacked form:

$$\mathbf{y}_t = X_t \boldsymbol{\theta} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_t),$$

where $X_t = (\tilde{X}_t, W_t)$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')$ is of dimension $k_\theta = k_\beta + k_\gamma$. If we assume $p = 3$ lags, then $k_\theta = 2592$ parameters to estimate.

In order to complete the model specification of the VAR(p) with stochastic volatility we need to include the time-varying volatility in the model. Thus, we include the log-volatilities $\boldsymbol{h}_t = (h_{1t}, \ldots, h_{Ht})$ for $t = 1, \ldots, T$. Following Cogley and Sargent (2005), we assume that the latent log-volatilities $\boldsymbol{h}_t$ evolve according to a random walk process

$$\boldsymbol{h}_t = \boldsymbol{h}_{t-1} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \Omega),$$

where $\boldsymbol{u}_t$ is a vector of i.i.d. residuals, $\Omega = \mathrm{diag}(\sigma^2_{h_1}, \ldots, \sigma^2_{h_H})$ and $\boldsymbol{h}_0$ is treated as a parameter to be estimated.

As stated in Cross et al. (2020), the prior specifications for the state variance, Ω, and for the initial state, $\boldsymbol{h}_0$, follow an independent prior distribution such as

$$\boldsymbol{h}_0 \sim \mathcal{N}(\boldsymbol{a_h}, V_h), \quad \sigma^2_{h_i} \sim \mathcal{IG}(\nu_{h_i}, S_{h_i}) \quad \text{for } i = 1, \ldots, H,$$

where $\mathcal{IG}(\cdot, \cdot)$ denotes an inverse Gamma distribution. Regarding the hyperparameters, we set $\boldsymbol{a_h} = \mathbf{0}$, $V_h = 10 \times I_H$; $\nu_{h_i} = 10$ and $S_{h_i} = 0.1^2(\nu_{h_i} - 1)$.

Regarding the prior distribution of the vectorized matrix of coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we assume independent Normal prior[3] specification of the form:

$$\boldsymbol{\beta} \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_\beta, \underline{V}_\beta); \quad \boldsymbol{\gamma} \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_\gamma, \underline{V}_\gamma),$$

where $\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\mu}}_\gamma$ are the prior means and $\underline{V}_\beta, \underline{V}_\gamma$ are the prior covariance matrix.

We can specify Eq. 2 in a different form by using a different representation of the $\tilde{X}_t$ matrix. In particular, we include some exogenous variables in the VAR, which leads to a VARX specification with stochastic volatility. The exogenous variables included in the analysis refer to both the demand and supply curves. As far as the former is concerned, we include the forecasted hourly demand $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Ht})'$ which contains variability around the expected levels of demand. As far as the supply is concerned, we consider fossil fuel prices, which, however, do not change over the 24 hours and are determined over the previous day, and the variability induced by the forecasted values for RES. Then, to summarize, we have included $\boldsymbol{w}_t = (w_{1t}, \ldots, w_{Ht})'$ for forecasted wind generation; $\boldsymbol{z}_t = (z_{1t}, \ldots, z_{Ht})'$ for forecasted solar power generation; and m_{t-1}, g_{t-1} and c_{t-1} for CO_2, gas and coal prices determined on the previous day, respectively.

From Eq 2, we redefine the matrix $\tilde{X}_t = I_H \otimes (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p}, \mathbf{d}'_t, \mathbf{x}'_t, \mathbf{w}'_t, \mathbf{z}'_t, m_{t-1}, g_{t-1}, c_{t-1})$, thus transforming Eq. 2 into a VARX(p)-SV, where $k_\beta = (Hp + K + 3H + 3)H$ is the dimension of the vector of coefficients $\boldsymbol{\beta}$.

In this paper, we consider a different specification of the time-varying covariance matrix that augments the (random walk) stochastic volatility specification to include fat tails. Thus, we

[3]Please note that we have also used a shrinkage prior of the Dirichlet-Laplace form for all the components of the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in both VAR and VARX specifications. However, the results were similar to those with the Normal prior, hence they have been omitted.

introduce a VAR(p) with stochastic volatility with Student-t error term and we introduce the degrees of freedom of a Student-t distribution. Differently from a strand of the literature (see Chan, 2020a,b), we have decided to include the degrees of freedom that change with the variables, such as electricity demand and solar power generation, which show higher values during daytime. Thus, $\nu_j > 0$ depends on $j = 1, \ldots, H$ to account for the varying dynamics across the hours.

The model specification follows Eq. (2) but the variance matrix Σ_t has a novel component. The VAR(p)-tSV is given by

$$\begin{aligned}
\boldsymbol{y}_t &= \tilde{X}_t \boldsymbol{\beta} + W_t \boldsymbol{\gamma} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(\mathbf{0}, \mathrm{diag}(\exp(h_{1t})/\lambda_{1t}, \ldots, \exp(h_{Ht})/\lambda_{Ht})), \\
\boldsymbol{h}_t &= \boldsymbol{h}_{t-1} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \mathrm{diag}(\sigma^2_{h_1}, \ldots, \sigma^2_{h_H})),
\end{aligned} \tag{3}$$

where $\lambda_{jt} \sim \mathcal{IG}(\nu_j/2, \nu_j/2)$ for every t and $j = 1, \ldots, H$. We consider two different options on the degree of freedom of the fat tail component. In the first one, the degrees of freedom ν_j are parameters to be estimated for each variable, thus we assume a Gamma prior for ν_j of the form $\nu_j \sim \mathcal{G}a(\underline{a}_\nu, \underline{b}_\nu)$; while in the second case, we fix the degrees of freedom $\nu_j = 5$ to ensure fat tails.

As for the Gaussian stochastic volatility model, we consider a specification with exogenous variables also for the fat tails case. Thus, we have a different specification of $\tilde{X}_t = I_H \otimes (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p}, \mathbf{d}'_t, \mathbf{x}'_t, \mathbf{w}'_t, \mathbf{z}'_t, m_{t-1}, g_{t-1}, c_{t-1})$ and we can define a VARX(p)-tSV model in the in-sample and forecasting exercise.

As anticipated, we do not encounter any ordering problems in the lower triangular matrix A_0 as instead in Carriero et al. (2019); Cross et al. (2020). In fact, in our analysis, we have a strong time dependence due to hourly specification and the variable ordering in the VAR-SV and VAR-tSV model will not affect the in-sample analysis and the relative forecasting performance of the models. We remember that the 24 hourly prices are set jointly the day before delivery, this implies that hourly prices cannot be treated iteratively across the day, but they have to be modelled as published by the system operator. This is a technical market requirement to plan the functioning of the system; although some adjustments can be undertaken only in subsequent market sessions, like the intra-daily and balancing ones, which close after the day-ahead session and determine, however, price series differing in their pricing mechanisms.

For comparison, we consider also a constant variance model, thus a VAR(p) and a VARX(p), where the error term is distributed with a constant covariance matrix $\Sigma_t = \Sigma$. For the prior assumption, we use the usual Bayesian prior specification for the covariance matrix and we consider an inverse Wishart distribution, $\mathcal{IW}(\underline{\nu}_0, \underline{\Psi}_0)$, where $\underline{\nu}_0$ are the prior degrees of freedom and $\underline{\Psi}_0$ is the prior scale matrix. A summary of the models used is presented in Table 1.

Models	Description
VAR	with Gaussian constant volatility
VARX	with Gaussian constant volatility
VAR-SV	with Gaussian stochastic volatility
VARX-SV	with Gaussian stochastic volatility
VAR-tSV	with stochastic volatility and t innovations (estimated ν_j)
VARX-tSV	with stochastic volatility and t innovations (estimated ν_j)
VAR-tSV$_{\boldsymbol{\nu}}$	with stochastic volatility and t innovations (fixed $\boldsymbol{\nu} = \nu_j = 5$)
VARX-tSV$_{\boldsymbol{\nu}}$	with stochastic volatility and t innovations (fixed $\boldsymbol{\nu} = \nu_j = 5$)

Table 1: List of Competing Bayesian VAR Models.

3 Data Description

Given the high penetration of renewables, as for wind in Germany and solar in Italy, we focus on these two countries and use hourly day-ahead auction prices directly collected from the corresponding power exchanges: the *European Energy Exchange* EEX for the former market, and the *Gestore dei Mercati Energetici* GME for the latter one, considering specifically the Italian single national prices PUN.

In addition, we consider the forecasted renewable generation from wind and solar power and the forecasted demand, together with closing settlement prices for fossil fuels to account for marginal production costs. As far as the forecasts for wind, solar and demand are concerned, they were provided by Thomson Reuters at hourly frequency. In details, we used values forecasted by the *operational* weather model provided by the *European Centre for Medium-Range Weather Forecast*, EC. This model runs at midnight and updates from 05.40 a.m. to 06.55 a.m., hence providing the latest information available to market operators to prepare their bidding strategy to be submitted into the day-ahead market by noon. In case of missing or unavailable (Italian) forecasts, we adopted this strategy to reconstruct the full required series. When the forecasts from the EC model running at midnight (its acronym is ECop00) were not available, we replaced missing observations according to the time of publication. In details, we considered secondly the forecasts provided by the *Global Forecast System* (GFS) running an ensemble model at midnight (that is GFSen00), alternatively in case of further missing observations we used the results of the operational model that was still running at midnight (that is GFSop00), and, if necessary, we used the same replacement scheme using respectively GFSen18, GFSop18, ECen12, according to the time of publication of their results. At the end of the process, residual missing observations were replaced with interpolated values. The two weather models differ in terms of randomness and resolution. The *operational* model is deterministic with high resolution and no involved randomness, whereas the *ensemble* model is a probabilistic model with lower resolution but with random variations of the initial weather conditions. Therefore, the latter simulates more weather instability by considering different weather scenarios.

Moving to fuel prices, we used the closing settlement prices for Coal ICE API2 CIF ARA (LMCYSPT), one month forward ICE UK natural gas prices (NATBGAS), and EEX-EU CO_2 Emissions E/EUA (EEXEUAS) converted into Euro/MWh, using the WMR&DS exchange rates for US$ to Euro (USEURSP) and GBP to Euro (UKEURSP). These data have been collected from Datastream and interpolated for missing quotations over weekends.

Finally, following Gianfreda et al. (2020), we have pre-processed series for solar power and used monthly and weekend dummy variables for calendar and weekly seasonality, respectively, with the latter ones containing also the indication for holidays. To summarize, we use hourly data for prices, forecasted demand, wind and solar PV generation, together with repeated daily data (across the 24 hours) for fossil fuel prices from 01 January 2016 to 31 December 2019.

4 Results

In this section, we illustrate the performance of the proposed Bayesian VAR models with different volatility structures.We first present results based on a full sample estimation and show that the introduction of stochastic volatility fits the data substantially better. Then, in a recursive out-of-sample forecasting exercise, we compare the forecast performance of the Bayesian VARs models by using different point and density measures.

In order to explain the model specification, we need to explain the lag structure p for our models and we follow common practice in the literature (Knittel and Roberts, 2005; Weron and Misiorek, 2008; Raviv et al., 2015; Gianfreda et al., 2020) and restrict lags to $t-1$, $t-2$, and $t-7$, which correspond to the previous day, two days before, and one week before the delivery time, recalling first similar conditions that may have characterized the market over the same hours and similar days (such as congestions and blackouts) and secondly the demand level during the days of the week. Hence, hourly prices with a reduced 7-lag structure are considered, and, with an abuse of notation in the remainder of the article, p = 3 is used to denote the maximum number of lags.

4.1 Estimation Results

In this section, we first present the empirical results for all the Bayesian VAR models, as listed in Table 1, which were obtained using the full sample from January 2016 to December 2019. We have also considered sub-samples for analysing the volatility in some periods, thus we reduce the sample from January 2016 to December 2017 and from January 2018 to December 2019. As an additional robustness check, we have performed our analysis on a yearly base, thus considering each year separately. The results presented in the Supplementary Material show that the volatility did indeed increase during the years 2018-2019.[4]

Results are very similar across all the models despite differences in the covariance structure. As described in Section 2, the stochastic volatility representation changes across hours, that is to say that we have it time-varying over the hours. Thus, in Figure 1 we depict the posterior means of the stochastic volatility, expressed in standard deviations $\exp(h_{kt}/2)$ for $k = 1, \ldots, H$ and $t = 1, \ldots, T$. We show the model comparisons over all 24 hours of the VARX model with constant volatility computed over 1 year centred rolling window, the VARX model with stochastic volatility and the VARX with fat tail stochastic volatility. First note that we left the same scale for emphasizing the difference across models. The volatility estimates from the VARX model with constant volatility are on average higher than for the time-varying volatility models. Figure 2 reports for each hour the number of times over our sample that the posterior means of the VARX model are higher than those of the VARX model with stochastic volatility and the VARX with fat tail stochastic volatility, respectively.[5] The posterior means of the VARX model are on average higher than those of VARX with fat tail stochastic volatility for all hours and for most of the hours excluding early hours at night when compared to the VARX with stochastic volatility. The average of this statistics over the 24 hours for German price is 64% in the VARX vs VARX-SV comparison and 89% in the VARX vs VARX-tSV comparison. For Italy, the results are qualitatively similar,

[4]For the sake of exposition, the paper reports all figures for the Germany example and the Supplementary Material shows the same results for Italy.

[5]When computing these statistics, we discard the initial 6-month data and the final 6-month data in order to compute rolling volatility from the VARX model.

with a 24-hour average of 60% for the VARX vs VARX-SV comparison and 87% for the VARX vs VARX-tSV comparison.

Moreover, it is interesting to observe that there are substantial changes across the hours of the days for the latter models and, in particular, looking at the fat tail representation of the error, the volatility shows high values during winters and lower values during summers recalling the calendar seasonality for all the four studied years. The VARX model with stochastic volatility seems to be more prone to spikes at different hours and different periods of the year. In general, the assumption of constant volatility could therefore imprecisely estimate the time-varying pattern of volatility.

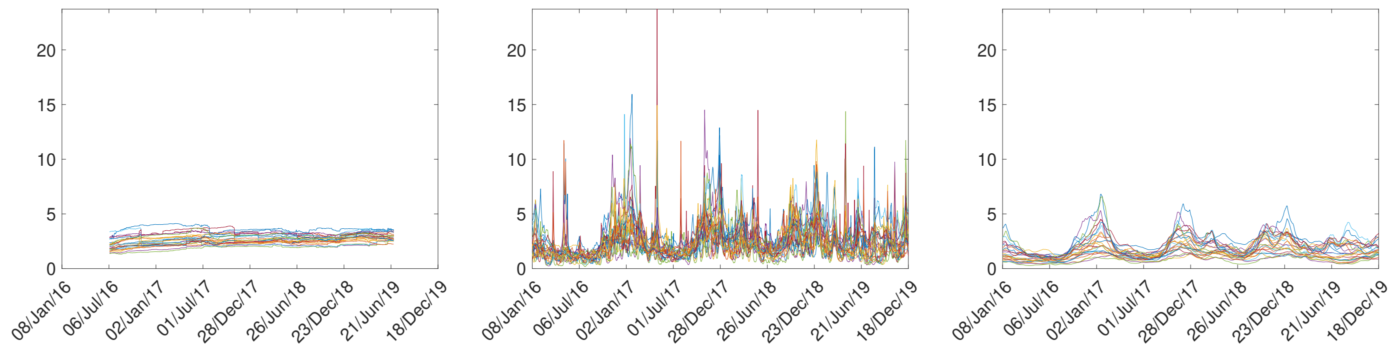

Figure 1: Posterior Means of the Stochastic Volatility Models, in standard deviations $\exp(h_t/2)$, observed in Germany. Model comparisons over all the 24 hours between the VARX computed over 1 year centred rolling window (left); the VARX-SV (middle) and the VARX-tSV (right).

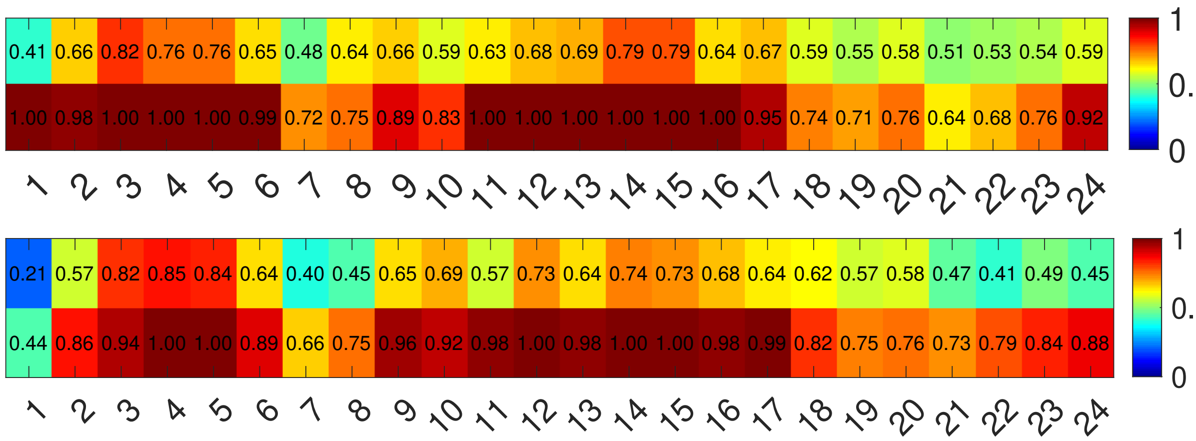

Figure 2: Number of times over the adjusted full sample that the posterior means of the VARX model are higher than those of the VARX model with stochastic volatility and the VARX with fat tail stochastic volatility. Top panel refers to Germany; bottom panel refers to Italy. First row for VARX vs VARX-SV and second rows for VARX vs VARX-tSV.

In addition, in Figure 3, we focus on the posterior means of the three specifications - VARX [6],

[6]It has been computed over 1 year centred rolling window, but we have also tested the usage of a 2 years centred rolling window with similar results; thus, they have been omitted but are available upon request.

VARX-SV and VARX-tSV - over the same hour, and we look in greater details at hours 10 (left), 14 (middle) and 18 (right), that is when RES show their higher production in connection with high levels of demand. Results for the other hours are provided in the Supplementary Material. Indeed, the model comparisons across the same hour emphasise indeed that the posterior means of the volatility estimated with a constant volatility structure (VARX) are higher than those with a time-varying structure (that is VARX-SV and VARX-tSV), and also that introducing fat tails in the error term (VARX-tSV in grey lines) systematically reduces the volatility over the sample considered and reductions are substantial especially during winters with this phenomenon being more clearly visible at hours 10 and 18. Surprisingly and more interestingly, the highest reductions are found at 14 when instead the economics of the energy systems suggests the major uncertainty due to the combination of forecasts errors for demand with those for wind and, especially, solar generations. This may also be the reason for which overall volatility increased over the last years of sample in hours 10 and 18, but less so for hour 14.

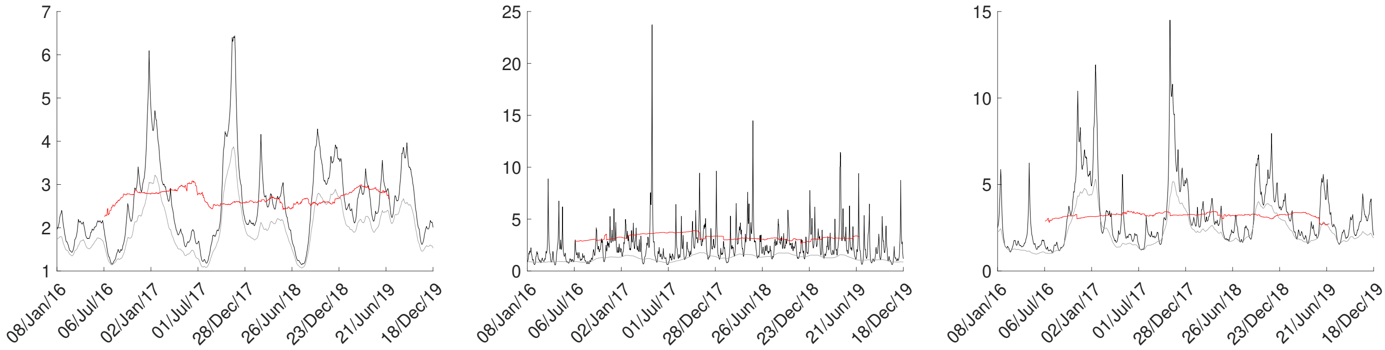

Figure 3: Posterior Means of the Stochastic Volatility Models, in standard deviations $\exp(h_t/2)$, observed in Germany. Model comparisons across hours 10 (left), 14 (middle) and 18 (right) for VARX (red line), VARX-SV (black line) and VARX-tSV (grey line).

Thus, to further inspect the effect of fat tails, we have investigated the differences between estimating or fixing the degrees of freedom at the same value over all the 24 hours. For the same three hours, Figure 4 shows the values for estimated $\boldsymbol{\nu}$ (black lines) and for $\boldsymbol{\nu}$ fixed to 3 (grey dashed lines) and to 5 (red lines). While black and red lines for the $\boldsymbol{\nu}$ estimated and fixed to 5 are almost perfectly overlapping at 10 and 18 (and also at 7-9 and 19-23 but less for $\boldsymbol{\nu}$ fixed to 3 at hours 12-16), the three dynamics decouple substantially at hour 14 when high volatility is observed for the degrees of freedom fixed to 5 and low values occur for estimated degrees of freedom. This then suggests how important is accounting for time varying changes estimated accounting for the characteristics of the 'current' sample, especially for hours with greater variability. These results confirm what was already suggested by Gianfreda and Bunn (2018) concerning the time-varying shape and tail dynamics exhibited by hourly electricity prices, even if they were considered individually each hour.

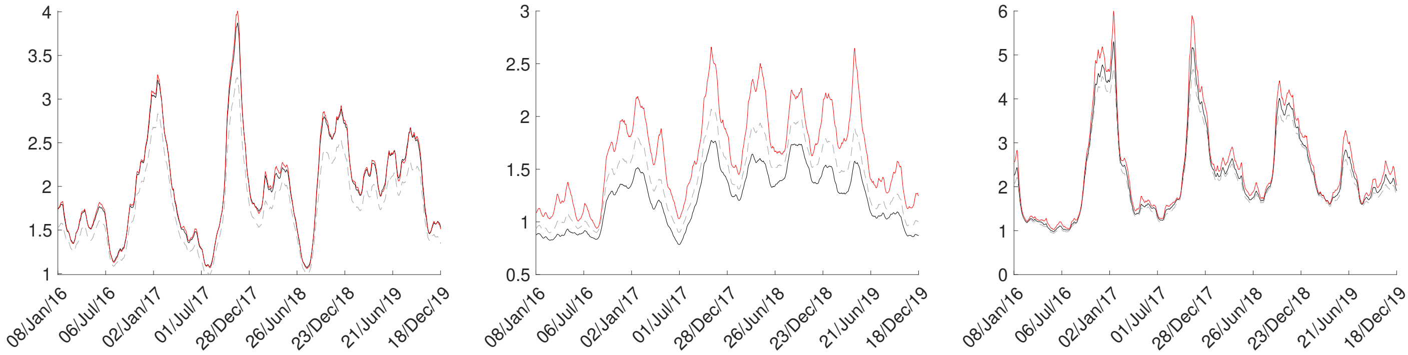

Figure 4: Posterior Means of the Stochastic Volatility Models VARX-tSV - in standard deviations $\exp(h_t/2)$ - with estimated $\boldsymbol{\nu}$ (black line), fixed $\boldsymbol{\nu} = 5$ (red line) and also $\boldsymbol{\nu} = 3$ (grey dashed) at hours 10 (left), 14 (centre) and 18 (right) in Germany.

Going into further details and looking at the distribution of degrees of freedom $\boldsymbol{\nu}$, Figure 5 shows the posterior means and distribution of ν over the 24 hours. The credibility intervals provide the uncertainty about estimated values and clearly show that it is high for hours 7-11 and 18-23 and instead it is low over hours 12-17. Reading these results together with those reflecting the differences in using estimated or fixed degrees of freedom, we can state that estimating or fixing the degrees of freedom does not implies dramatic changes at hours 10 and 18 even if the uncertainty is observed to be high; hence, practitioners can choose the approach they prefer. On the contrary, attention must be paid on central hours since, for instance, the low uncertainty at hour 14 is coupled with substantial differences between estimated or ex-ante fixed values, and in the latter case the choice of fixing degrees of freedom may produce substantial overestimation. Then, a dynamic estimation is suggested for these central hours, since there is a strong influence of the fat tailness of the data. Same results are found for Italy, and these have been reported in the Supplementary Material.

Given the large size of implemented models, it is worth noticing the computational cost in the estimation process. Table 2 shows the timing expressed in seconds to obtain $5,000$ posterior draws for the models considered in our analysis and reported in Table 1. The estimation period is the full sample ranging from January 2016 to December 2019 and the models are estimated using Matlab 2019b on a Macbook Pro with an Intel Core i7 @2.70 GHz processor and 16 GB memory. Computational times more than double when adding X variables and, more importantly,

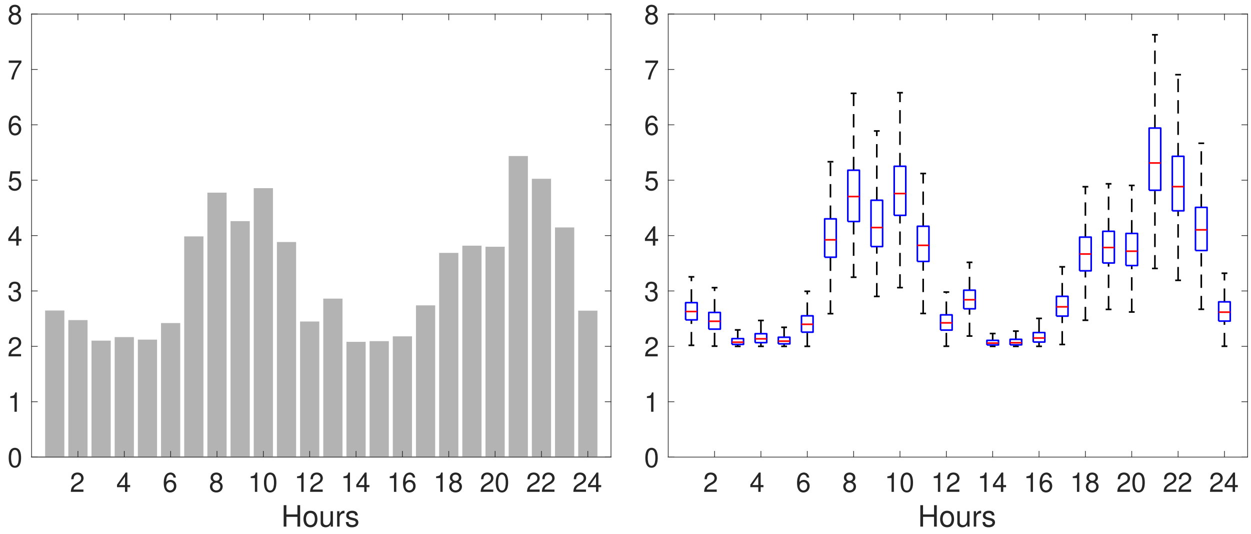

Figure 5: Posterior Means (left) and Posterior Distribution (right) of the Degrees of Freedom ν across the 24 Hours for the VARX-tSV Model for Germany. The red horizontal line represents the posterior mean, whereas the blue box indicates the 75% credibility interval.

increase by multiples when adding stochastic volatility and less dramatically when also adding fat tails. However, the estimation time for the most computational intense model is always below eight minutes.

	Const	SV	tSV	tSV ($\boldsymbol{\nu} = 5$)
Germany				
VAR	60.7	346.5	391.5	357.1
VARX	131.1	421.6	453.3	450.6
Italy				
VAR	61.2	358.4	388.6	385
VARX	133.8	439.9	463.5	460.3

Table 2: The computational times (in seconds) to obtain 5,000 posterior draws under each model specification on the full sample size ($T = 1454$).

4.2 Forecast Results

4.2.1 Forecast assessment

In order to assess the goodness of our forecasts, the rest of this section will focus on describing the point and density metrics. As a point forecast measure, we apply the root-mean-square errors (RMSEs) for each of the hourly prices, as well as the RMSEs for the daily average and of an average restricted only to the central hours. The RMSE for $h = 1, \ldots, 24$ hourly prices is calculated as

$$\text{RMSE}_h = \sqrt{\frac{1}{T-R}\sum_{t=R}^{T-1}\left(\hat{y}_{h,t+1|t} - y_{h,t+1}\right)^2},$$

where T is the number of observations, R is the length of the rolling window and $\hat{y}_{h,t+1|t}$ are the individual hourly price forecasts. In addition, we analyse the average RMSEs on all the 24 hours (RMSE_{Avg}) and on the hours from 8 a.m. to 8 p.m. (peak hours, $\text{RMSE}^P_{\text{Avg}}$), calculated as follows:

$$\text{RMSE}_{\text{Avg}} = \frac{1}{24}\sum_{h=1}^{24}\text{RMSE}_h; \quad \text{RMSE}^P_{\text{Avg}} = \frac{1}{13}\sum_{h=8}^{20}\text{RMSE}_h. \tag{4}$$

On the other hand, we evaluate density forecasts by using average continuous ranked probability score (CRPS) and the quantile CRPS (Gneiting and Raftery, 2007; Gneiting and Ranjan, 2011). These measures have advantages over the log score, in particular, they do a better job of rewarding values from the predictive density that are close to - but not equal to - the outcome, and they are less sensitive to outlier outcomes.

The CRPS, defined such that a lower number is a better score, is given by

$$\text{CRPS}_{h,t}(y_{h,t+1}) = \int_{-\infty}^{\infty}\left(F(z) - \mathbb{I}\{y_{h,t+1} \leq z\}\right)^2 dz = E_f|Y_{h,t+1} - y_{h,t+1}| - 0.5E_f|Y_{h,t+1} - Y'_{h,t+1}|, \tag{5}$$

where F denotes the cumulative distribution function associated with the predictive density f, $\mathbb{I}\{y_{h,t+1} \leq z\}$ denotes an indicator function taking the value 1 if $y_{h,t+1} \leq z$ and 0 otherwise, and $Y_{h,t+1}$ and $Y'_{h,t+1}$ are independent random draws from the posterior predictive density. In the same way we can construct the average CRPS over the 24 hours and over peak hours on day $t+1$.

Regarding the quantile CRPS, the quantile-weighted versions of the continuous ranked probability score is defined as:

$$S_{h,t}(y_{h,t+1}) = \int_0^1 \text{QS}_\alpha\left(F^{-1}(\alpha), y_{h,t+1}\right)\omega(\alpha)d\alpha, \tag{6}$$

where $\text{QS}_\alpha\left(F^{-1}(\alpha), y_{h,t+1}\right)$ is the quantile score defined as

$$\text{QS}_\alpha\left(F^{-1}(\alpha), y_{h,t+1}\right) = 2\left(\mathbb{I}\{y_{h,t+1} \leq F^{-1}(\alpha)\} - \alpha\right)\left(F^{-1}(\alpha) - y_{h,t+1}\right)$$

with $F^{-1}(\alpha)$ the quantile forecast and $\alpha \in (0,1)$. When $\omega(\alpha) = 1$, we have an uniform weight, thus an unweighted continuous ranked probability score.

The nonnegative weight function on the unit interval $\omega(\alpha)$ can take on different specifications in order to assess centre or tails emphasis. In particular, $\omega(\alpha) = \alpha(1-\alpha)$ defines a centre emphasis, while a tails emphasis is computed from $\omega(\alpha) = (2\alpha - 1)^2$. The right and left emphasis are denoted as $\omega(\alpha) = \alpha^2$ and $\omega(\alpha) = (1-\alpha)^2$, respectively. As for the RMSE and CRPS, we can compute the average quantile CRPS over the 24 hours and over the peak hours on day $t+1$. In what follows, we indicate the average center quantile CRPS with CQ-CRPS, the average right tails quantile CRPS with RQ-CRPS, and the average left tails quantile CRPS with LQ-CRPS, respectively.

Moreover, in this paper we report the RMSEs, average CRPS and average quantile CRPSs for the baseline VAR model with constant volatility and for every third hour[7]. For the other VAR models, we report the ratios computed between the RMSE of the current models and the RMSE of the baseline VAR model. Then, entries of less than 1 indicate that the given current model yields forecasts more accurate than are those provided from the baseline, and similarly for the CRPS and the quantile CRPS.

In addition, to provide a rough gauge of whether the differences in forecast accuracy are significant, we apply Diebold and Mariano (1995) t-tests for equality of the average loss to compare predictions of alternative models to the benchmark for a given horizon h[8]. The differences in accuracy that are statistically different from zero are denoted with one, two, or three asterisks, corresponding to significance levels of 10%, 5%, and 1%, respectively. The underlying p-values are based on t-statistics computed with a serial correlation-robust variance, using the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992). Our use of the Diebold-Mariano test, with forecasts from models that are, in many cases, nested, is a deliberate choice, as in Clark and Ravazzolo (2015), and, as noted by Clark and West (2007) and Clark and McCracken (2012), this test is conservative and might result in under-rejection of the null hypothesis of equal predictability. We report p-values based on one-sided tests, taking the VAR as the null and the other current models as the alternative.

Finally, the Model Confidence Set procedure of Hansen et al. (2011) across models for a fixed horizon have been employed to jointly compare their predictive power without disentangling

[7]Tables with all hours and all models are available in the Supplementary Material.

[8]In our application for testing density forecasts, we use equal weights without adopting a weighting scheme, as in Amisano and Giacomini (2007).

between constant or time-varying volatility. The R package MCS detailed in Bernardi and Catania (2016) has been used, and the differences have been tested separately for each hour and model, repeating the full process for both countries. Results are discussed in the following section.

4.2.2 Forecasting Results

Our results are based on a one-step ahead rolling forecasting process with a window size of two years for both countries. Please note that the initial estimation sample goes from 1 January 2016 to 31 December 2017 and then the forecasting evaluation period starts on 1 January 2018 and ends on 31 December 2019; hence, there is a total of 731 in-sample forecasts.

Results refer to the performances of our different multivariate models from the simplest one (with lags and dummy variable, the benchmark VAR) to more complex ones including constant or different time-varying volatility specifications as well as fundamental drivers (as fuel prices and forecasts for demand and renewable energy sources). Tables 3 and 4 report the results across Germany and Italy respectively, every third hour. Whereas, results across all 24 hours and extensive model comparisons are shown in Tables S.1-S.12 in the Supplementary Material.

As expected, the forecasting performance decreases across the peak hours (that is between hour 8 and hour 20) and this is consistent with the high uncertainty affecting demand levels during daytime and also supply; and this is consistent across all electricity markets. Hence, the benchmark VAR models show the highest RMSEs at hours 14-16 in Germany (around 13€/MWh) and at peak hours 9-10 & 15-20 in Italy (around 8€/MWh). These differences can be explained by the intra-daily dynamics of forecasted demand and RES generation, which differ substantially across countries because of the diverse geographical conditions affecting, for instance, hours of daylight, solar radiation and wind speed. Further details and comparisons of the intra-daily dynamics of demand and RES generation are presented in Gianfreda et al. (2020) and Gianfreda et al. (2016).

The most interesting aspect addressed here is the expected forecasting improvement resulting from accounting for volatility together with fundamental drivers (demand, RES and fuels).

First of all, our results show that in both markets the Bayesian multivariate models with stochastic volatility (VARX-SV) exhibit substantial improvements with respect to the benchmark VARs with a constant volatility and the VARX. Including fundamental drivers in the modelling of price variability increase forecast accuracy and even larger gains are obtained by extending with time-varying volatility. The Model Confidence Set shows that the VAR and VARX models are never included in the model set when predicting German electricity prices in terms of RMSE and CRPS metrics and only at low volatile early-morning hours for Italian prices.

Further improvements can be observed when we include the Student-t stochastic volatility. Again, this occurs for both point and density metrics. Moreover, we observe no difference between estimating the degrees of freedom or fixing them equal to an ex-ante selected value (as for instance 5) across hours. This confirms our previous findings and proves that both ways produce similar forecasting results in Germany. Therefore, practitioners can make an indifferent decision regarding their modelling strategy, but we suggest estimating the degrees of freedom across hours in order to address different degrees of fat tails over the 24 hours. On the contrary, for Italy there is no empirical evidence of differences between Gaussian and fat tail stochastic volatility models across the considered forecasting measures. On one hand, this is in line with the lower bound market constraint applied to Italian prices; on the other hand, this may be due to the different levels of RES penetration observed in Italy, especially wind when compared to its levels observed in Germany.

Looking specifically at the metrics, the average reductions in loss function are similar for both metrics of about 20% in Germany and 5% in Italy. Forecasting gains in terms of the RMSE increase from hour 8 to the end of the day, as shown more clearly from the last column of Tables 3 and 4 in which, however, only central hours are considered. More generally, the RMSEs decrease from the VARX to the VARX-tSV across all hours and more in Germany than in Italy. Similarly for the CRPSs, for which we observe reductions of almost 10% in Germany and 5% in Italy. In Germany the use of fat tails error (Student t distribution) leads to improvements of 20% in both metrics, while we do not observe similar reductions in Italy.

Considering the tails, results in terms of the average LQ-CRPS and RQ-CRPS confirm the expected lower gains in Italy, of about 0.8 with respect to the benchmark levels for both left and right tails. Surprisingly, we observe substantial forecasting improvements of, however, equal magnitude on both left and right tails in Germany, of about 0.6 with respect to the benchmark levels. These details are reported in the Supplementary Material for the On-line Appendix. To summarize them, we have displayed the LQ-CRPS in Table 3, in order to capture the negative prices in Germany; and the RQ-CRPS in Table 4, in order to capture more positive prices in Italy.

As argued by Gianfreda and Bunn (2018) in their analysis on German individual hours selected according to intra-daily profiles, wind and solar generation reduce the skewness of hourly electricity prices, with this phenomenon being more evident at hours 12-13 because solar is at its maximum level. Moreover, they add that both increase the kurtosis of electricity prices at peak hour 19. On the contrary, hour 3 shows higher volatility because more negative price spikes are observed compared to hours 12-13 and 19.

Therefore, considering Germany and its off-peak1 prices (that is in the early morning from hour 1 to 7), lower quantiles are of most practical interest for the occurrence of down spikes. Then, looking for instance at hour 3 we observe that the LQ-CRPS shows limited improvements when considering constant or time-varying with Student-t volatility. This may be due to lower values observed during off-peak1 hours than those observed in peak (8-20) and off-peak2 (21-24) hours: there is indeed a jump in this metric from 0.413 at hour 1 to 1.158 at hour 15, which, however, still confirms the convenience in implementing more complex models with regressors and time-varying volatility with estimated or fixed degrees of freedom. Instead, for hours 12-13, both the high and low quantiles are of interest and here we observe substantial and significant improvements in both the LQ-CRPS and RQ-CRPS from the VARX-tSV when indifferently estimating or fixing ν. Similar comments apply at hours 19-21 when the high quantiles are the most interesting for the risk of high prices, because RES decreases but demand is still at high levels.

These comments, however, do not apply to the Italian prices since they are affected by RES with a lower magnitude compared to Germany and, more importantly, they are not allowed to become negative. However, it is still interesting to observe that during peak hours the inclusion of time-varying volatility and regressors improves both point and density metrics.

Table 3: Forecasting Metrics for Germany. Note that real values are used for the Bayesian VAR model, and ratios for all the other models

Hour	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
RMSE										
VAR	5.003	6.792	9.070	11.288	11.908	13.286	11.675	10.130	8.070	10.704
VARX	1.099	0.939	0.875**	0.784***	0.763***	0.759***	0.747***	0.674***	0.739	0.718
VARX-SV	1.050	0.892***	0.798***	0.718***	0.698***	0.711***	0.689***	0.677***	0.675	0.658
VARX-tSV	1.036	0.886***	0.786***	0.704***	0.680***	0.712***	0.684***	0.671***	0.670	0.651
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.040	0.882***	0.785***	0.707***	0.678***	0.706***	0.684***	0.676***	0.667	0.650
CRPS										
VAR	2.545	3.416	4.502	5.740	6.108	6.635	6.075	5.294	4.093	5.485
VARX	1.101	0.960*	0.906***	0.784***	0.753***	0.749***	0.742***	0.679***	0.740	0.716
VARX-SV	1.035	0.871***	0.791***	0.695***	0.686***	0.677***	0.660***	0.642***	0.652	0.634
VARX-tSV	0.996	0.848***	0.763***	0.675***	0.641***	0.648***	0.652***	0.641***	0.633	0.614
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.001	0.849***	0.761***	0.677***	0.643***	0.649***	0.651***	0.639***	0.632	0.614
CQ-CRPS										
VAR	0.305	0.412	0.540	0.682	0.718	0.794	0.715	0.619	0.487	0.650
VARX	1.097	0.932**	0.878***	0.772***	0.741***	0.743***	0.734***	0.673***	0.734	0.715
VARX-SV	1.020	0.850***	0.785***	0.683***	0.682***	0.685***	0.653***	0.636***	0.645	0.630
VARX-tSV	0.983	0.837***	0.761***	0.667***	0.639***	0.654***	0.647***	0.635***	0.629	0.614
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.986	0.833***	0.755***	0.667***	0.637***	0.651***	0.643***	0.633***	0.625	0.611
LQ-CRPS										
VAR	0.413	0.577	0.722	0.908	0.989	1.101	0.961	0.875	0.665	0.881
VARX	1.045	0.911***	0.889***	0.774***	0.749***	0.730***	0.731***	0.657***	0.724	0.707
VARX-SV	1.028	0.849***	0.795***	0.690***	0.683***	0.680***	0.646***	0.625***	0.652	0.633
VARX-tSV	1.001	0.851***	0.778***	0.675***	0.648***	0.655***	0.639***	0.625***	0.642	0.619
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.001	0.845***	0.775***	0.677***	0.649***	0.654***	0.635***	0.622***	0.638	0.617

Notes:

[1] Please refer to Section 2 for the details on models. The 'X' indicates models with exogenous variables. All forecasts are produced with one-step-ahead rolling window process.

[2] ***, ** and * indicate that ratios are significantly different from 1 at 1%, 5% and 10%, according to the Diebold-Mariano test.

[3] Grey cells indicate models that belong to the Superior Set of Models delivered by MCS procedure at confidence level 10%.

[4] The results with a fixed $\boldsymbol{\nu} = 3$ are similar with results obtained with $\boldsymbol{\nu} = 5$, thus they have not been reported here but are available in the supplementary material.

Table 4: Forecasting Metrics for Italy. Note that real values are used for the Bayesian VAR model, and ratios for all the other models

Hour	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
RMSE										
VAR	4.431	5.269	5.806	8.448	6.872	8.479	8.568	6.289	5.054	6.902
VARX	1.030	1.021	0.983	0.905**	0.916***	0.890***	0.979	0.960*	0.953	0.900
VARX-SV	1.015	1.017	0.959	0.859***	0.893***	0.862***	0.899***	0.897***	0.906	0.856
VARX-tSV	1.004	1.022	0.960	0.871***	0.896***	0.869***	0.904***	0.894***	0.911	0.864
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.001	1.024	0.958	0.865***	0.894***	0.866***	0.901***	0.892***	0.908	0.860
CRPS										
VAR	2.446	2.925	3.073	4.498	3.652	4.513	4.607	3.427	2.733	3.721
VARX	1.037	1.019	0.995	0.912***	0.900***	0.889***	0.975	0.957**	0.960	0.902
VARX-SV	1.017	0.999	0.959**	0.851***	0.871***	0.844***	0.876***	0.870***	0.899	0.840
VARX-tSV	1.001	1.002	0.961**	0.856***	0.866***	0.847***	0.878***	0.867***	0.901	0.843
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.999	1.001	0.958**	0.850***	0.863***	0.844***	0.874***	0.864***	0.899	0.840
CQ-CRPS										
VAR	0.281	0.334	0.360	0.523	0.425	0.525	0.532	0.392	0.314	0.432
VARX	1.034	1.025	0.991	0.890***	0.899***	0.875***	0.970*	0.955**	0.953	0.891
VARX-SV	1.005	0.983	0.936***	0.832***	0.864***	0.836***	0.874***	0.860***	0.881	0.826
VARX-tSV	0.986	0.989	0.937***	0.841***	0.861***	0.845***	0.880***	0.858***	0.884	0.832
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.985	0.991	0.936***	0.833***	0.858***	0.839***	0.872***	0.854***	0.881	0.827
RQ-CRPS										
VAR	0.393	0.444	0.478	0.716	0.583	0.719	0.719	0.549	0.423	0.585
VARX	1.015	1.037	1.018	0.924***	0.920***	0.903***	0.985	0.965*	0.983	0.921
VARX-SV	1.005	1.013	0.981	0.869***	0.903***	0.866***	0.910***	0.880***	0.935	0.872
VARX-tSV	0.988	1.011	0.982	0.875***	0.899***	0.873***	0.908***	0.879***	0.939	0.874
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.985	1.013	0.980	0.870***	0.896***	0.871***	0.907***	0.878***	0.937	0.873

Notes: Please see the notes to Table 3.

5 Conclusions

Modelling day-ahead electricity prices has become extremely important for understanding the energy system and providing empirical support to policy makers in a context where uncertainty is progressively increasing as a consequence of changing weather conditions due to climate change. In this regard, appropriate models may also produce useful forecasts, which help market operators plan their generation schedules while accounting (and then adapting) to the imperfect predictability of both demand and RES generation. This is also extremely relevant for the transmission system operators who must guarantee the continuous balance between demand and supply.

In this framework, we address the less explored issue of multivariate models in which volatility dynamics are included. We have questioned whether the inclusion of a constant or a time-varying volatility structure can better detect the movements of electricity prices in two important European countries, namely Germany and Italy. Thus, we propose high dimensional VAR models with different stochastic volatility representations in which fundamental drivers are included as exogenous variables, these are forecasted demand, renewable energy sources and fuels. In particular, we assume that the time-varying volatility changes across hours and drives the dependence in a time-ordered structure, then we do not charge any ordering label problems as stated in the literature (Carriero et al., 2019; Cross et al., 2020).

Using only a lagged representation of the data or adding different exogenous variables, we find empirical supporting evidence for VAR models with stochastic volatility against the conventional VAR. Indeed, most of the gains appear to come from allowing stochastic volatility rather than constant volatility. In particular, during some hours of the day, the assumption of fat tails in the error term improves the detection of time-varying volatility. Furthermore, in a recursive forecasting exercise, we find that models with exogenous variables also show improvements in both point and density forecasts. In addition, the combined inclusion of time-varying volatility plus exogenous variables lead to even better point and density metrics.

For future research, it would be interesting to extend these models specification by including a global shrinkage prior for both exogenous and lagged variables or alternatively to use different priors in all the parameters.

References

Amisano, G. and Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25(2):177–190.

Andrews, D. and Monahan, J. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60(4):953–966.

Bauwens, L., Hafner, C. M., and Pierret, D. (2013). Multivariate volatility modeling of electricity futures. *Journal of Applied Econometrics*, 28(5):743–761.

Bernardi, M. and Catania, L. (2016). Portfolio Optimisation Under Flexible Dynamic Dependence Modelling. *ArXiv e-prints*.

Carriero, A., Clark, T. E., and Marcellino, M. (2016). Common Drifting Volatility in Large Bayesian VARs. *Journal of Business & Economic Statistics*, 34(3):375–390.

Carriero, A., Clark, T. E., and Marcellino, M. (2019). Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212(1):137–154.

Chan, J. C. C. (2020a). Large Bayesian VARs: A Flexible Kronecker Error Covariance Structure. *Journal of Business & Economic Statistics*, 38(1):68–79.

Chan, J. C. C. (2020b). *Large Bayesian Vector Autoregressions*, pages 95–125. Springer International Publishing, Cham.

Chan, J. C. C. and Eisenstat, E. (2018). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics*, 33(4):509–532.

Clark, T. E. and McCracken, M. W. (2012). Reality checks and comparisons of nested predictive models. *Journal of Business & Economic Statistics*, 30(1):53–66.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2):262–302.

Cross, J. L., Hou, C., and Poon, A. (2020). Macroeconomic forecasting with large Bayesian VARs: Global-local priors and the illusion of sparsity. *International Journal of Forecasting*, 36(3):899–915.

Damm, A., Köberl, J., Prettenthaler, F., Rogler, N., and Töglhofer, C. (2017). Impacts of +2°C global warming on electricity demand in Europe. *Climate Services*, 7:12 – 30. IMPACT2C - Quantifying projected impacts under 2 °C warming.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263.

Gianfreda, A. and Bunn, D. (2018). A stochastic latent moment model for electricity price formation. *Operations Research*, 66(5):1189–1203.

Gianfreda, A., Parisio, L., and Pelagatti, M. (2016). Revisiting long-run relations in power markets with high RES penetration. *Energy Policy*, 94:432 – 445.

Gianfreda, A., Ravazzolo, F., and Rossini, L. (2020). Comparing the forecasting performances of linear models for electricity prices with high RES penetration. *International Journal of Forecasting*, 36(3):974–986.

Gneiting, T. and Raftery, A. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of American Statistical Association*, 102(477):359–378.

Gneiting, T. and Ranjan, R. (2011). Comparing density forecasts using threshold- and quantile-weighted proper scoring rules. *Journal of Business and Economic Statistics*, 29(3):411–422.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, 79:453–497.

Huber, F. and Feldkircher, M. (2019). Adaptive Shrinkage in Bayesian Vector Autoregressive Models. *Journal of Business & Economic Statistics*, 37(1):27–39.

Huber, F., Koop, G., and Onorante, L. (2020). Inducing Sparsity and Shrinkage in Time-Varying Parameter Models. *Journal of Business & Economic Statistics*, pages 1–15.

Karakatsani, N. and Bunn, D. W. (2010). Fundamental and behavioural drivers of electricity price volatility. *Studies in Nonlinear Dynamics & Econometrics*, 14(4):1–42.

Knittel, C. and Roberts, M. (2005). An empirical examination of restructured electricity prices. *Energy Economics*, 27(5):791–817.

Koop, G., Korobilis, D., and Pettenuzzo, D. (2019). Bayesian compressed vector autoregressions. *Journal of Econometrics*, 210(1):135–154.

Raviv, E., Bouwman, K. E., and van Dijk, D. (2015). Forecasting day-ahead electricity prices: Utilizing hourly prices. *Energy Economics*, 50:227 – 239.

Weron, R. and Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4):744–763.

Centre for Applied Macroeconomics and Commodity Prices (CAMP)
will bring together economists working on applied macroeconomic issues, with special emphasis on petroleum economics.

BI Norwegian Business School
Centre for Applied Macro - Petroleum economics (CAMP)
N-0442 Oslo

www.bi.no/camp

CAMP Working Paper Series
ISSN: 1892-2198