L Kang, S Walker, P Damien and D W Bunn
Bayesian estimation of electricity price risk with a multi-factor mixture of densities
Article

Kang, L, Walker, S, Damien, P and Bunn, D W

(2022)

Bayesian estimation of electricity price risk with a multi-factor mixture of densities.

Quantitative Finance, 22 (8). pp. 1535-1544. ISSN 1469-7688

DOI: https://doi.org/10.1080/14697688.2022.2052165

Taylor & Francis (Routledge)
https://www.tandfonline.com/doi/abs/10.1080/146976...

Bayesian Estimation of Electricity Price Risk with a Multi-factor Mixture of Densities

Li Kang[*], Stephen Walker[†], Paul Damien[‡] and Derek Bunn[§]

Abstract

The risks in daily electricity prices are becoming substantial and it is clear that improvements in price density forecasting can translate into improved improved risk management. However, the specification of the most appropriate price density function is challenging as the best functional forms differ by time of day evolve over time, dynamically respond to fluctuating exogenous factors such as wind speed and solar irradiance. This research develops and tests a new flexible functional form based upon the Gamma Mixture of Uniform (GMU) densities which effectively avoids the choice of a particular density function and has conditional moments specified as a function of the dynamic exogenous drivers. Empirical testing shows that it outperforms the multi-factor skewed student-t family of densities, previously advocated in this context. Additionally, using Bayesian estimation the new methodology provides a complete description of the uncertainty in the estimation of the coefficients for those exogenous factors. Empirical testing on day-ahead hourly electricity prices in the German market from 2012 to 2016, where renewable energy sources, such as wind and solar, play a critical role in the formation of electricity price risk, validates the extra accuracy of this formulation.

Keywords: Bayesian inference; Electricity prices; Markov chain Monte Carlo; Merit-order effects; Stochastic skewness.

[*]Department of Statistics and Data Science, University of Texas in Austin; likang@utexas,edu

[†]Department of Mathematics, University of Texas in Austin; s.g.walker@math.utexas.edu

[‡]Corresponding author: McCombs School of Business, University of Texas in Austin; paul.damien@mccombs.utexas.edu

[§]London Business School; dbunn@london.edu

1 Introduction

Electricity markets now exhibit high trading risks. There is the rapid technological switch to renewable generation together with a much greater consumer engagement though solar panels, electric vehicles and smart energy management systems. The effect of these has been to increase uncertainties and hence price risk in the daily wholesale markets.This risk has been serious, and financial distress for both incumbent generators and new entrant retailers has followed. For example, in Britain, over 60 retailers competed for customers in 2017, but 19 of those failed to pay their trading charges in 2018 and had to exit the market, with even more (26) failing in 2021. In most cases this was due to inadequate hedging of price exposures[1]. More precise risk management evidently requires more precise estimations of the price density functions; see, for example, (Brusaferri et al., 2019) who used Bayesian deep learning to develop day-ahead probabilistic forecasts, whilst (Canelas et al., 2020) show the advantages of reducing trading costs and risk in the Iberian context with a more accurate forecasting specification. Beyond the electricity context, it is becoming widely recognized that density specifications for risk management could benefit from higher moment estimations, with skewness and kurtosis often having crucial implications. For example, four moments are used in asset portfolio optimizations (Jondeau and Rockinger, 2006; Giesecke et al., 2014) and in option pricing models (de Jong and Huisman, 2000; Borland and Bouchaud, 2004; Aboura and

[1]Insights: Defaults and SoLRs increased in 2018 by Emma Tribe in *Elexon Insights*, https://www.elexon.co.uk/operations-settlement/balancing-and-settlement/trading-charges/elexon-insights-defaults-solrs-increased-2018/)

Maillard, 2016), whilst the imperative of Value-at-Risk compliance has required accuracy at the extreme quantiles of price returns. In general, this higher moment requirement has led to the use of various four-parameter distributions with the skewed student-t becoming rather popular.

Electricity is a good example of this need for higher moment specifications and furthermore requires conditional price densities where exogenous factors, such as renewable generation, have explicit dynamic impacts on the price density function. Several electricity price applications have used densities whose first three or four moments have been functions of fundamental exogenous factors, which are estimated from a general additive model (Serinaldi, 2011; Gianfreda and Bunn, 2018; Abramova and Bunn, 2020); this is termed multi-factor models. Various density functions have been chosen in those cases. However, the choice of the appropriate density function is not straightforward with different densities performing best at different hours of the day and with their relative performances changing over time. Furthermore, as exogenous factors such as wind speed change, the densities can alter dramatically with skewness, for example, flipping from positive to negative under high winds. Our contribution is to resolve this density specification problem by developing a highly flexible formulation based upon a conditional gamma mixture of uniform densities (GMU) which can evolve, via dynamic estimation, to represent a wide range of densities as a function of the exogenous drivers (such as wind speed). We benchmark our approach against the best performing model in those previous electricity price applications, namely the skewed student-t family of densities, which the GMU outperforms in terms of accuracy of the predictive price densities. Furthermore, by using Bayesian estimation, we contribute by fully representing the uncertainty in the coefficients for those driving factors which may be useful in pricing partial hedging instruments such as parametric weather insurance (e.g. on wind speed).

This research is closest to Hagfors et al. (2016) who model the German day-ahead electricity prices via non-Bayesian models. Like us, they too provide useful insights on how renewable energy sources, specifically wind, impact the probability of negative prices and positive spikes. They show that extremely high and negative prices have different drivers. Moreover, the found that wind power is of particular relevance in relation to negative price occurrences. Another interesting study (Benth et al. (2020)) uses a multivariate Ornstein-Uhlenbeck process to hedge wind risk by modeling wind indexes. However, unlike our research, their focus is on finding optimal hedging strategies using exchange-traded wind power futures in a portfolio. Our contribution is with a new multi-factor price density forecasting model which has not previously been developed and which appears to outperform current best practice.

The GMU was introduced by Kang et al. (2020), based on a simple skew transformation that allows one to model varying levels of skewness and tail behaviors. Denoting the four-parameter family of GMU densities as $Y \sim \mathrm{GMU}(\mu, \alpha, \beta, \lambda)$, where μ, α, β and λ are the location, kurtosis, scale and skewness parameters, respectively, then, a member from this family has probability density function $f(y)$ given by:

$$f(y) = \frac{\beta \, \mathrm{sech}(\lambda)}{2(\alpha - 1)} \, \gamma\big\{ r(y - \mu, \lambda) \big\}, \tag{1}$$

where

$$\gamma(t) = \textstyle\int_t^\infty \Gamma(z|\alpha - 1, \beta) \, dz \quad \text{and} \quad r(\xi, \lambda) = \max\big\{ -\xi \exp(\lambda), \xi \exp(-\lambda) \big\}.$$

Kang et al. (2020), using various parameter combinations, show that the GMU is more flexible than the skewed student-t, the asymmetric Laplace distribution, etc. Moreover, they show that the GMU includes, among others, the Laplace and the normal distributions as special cases. While Kang et al. (2020) showed that the GMU density outperforms several alternatives, including the skewed-t, skewed-normal, asymmetric Laplace, etc., that research did not address how, in a time series context, the conditional moments of the GMU density could be estimated dynamically as functions of several exogenous factors. The extension to a multi-factor GMU is one of the main theoretical contributions in this paper, as well as demonstrating its applied benefits in the electricity price context. The second theoretical departure is that we implement a full Bayesian analysis for the resulting latent moment model. A Bayesian approach is useful in risk estimation, since we can obtain probability distributions for all the random parameters in the model, as well as the consequent predictive distributions. We use the electricity market setting, and German prices in particular, because that context has provided previous developments on this class of methods against which we wish to benchmark our multi-factor GMU approach. In this context, our GMU family of densities outperforms, in terms of density risk estimation, the alternative approaches to electricity price risk. We demonstrate this using the pinball loss function for the extreme quantiles.

In the next Section, we describe the practical context and the data. Section 3 introduces the GMU family of multi-factor moment models; a technical appendix detailing the MCMC algorithm is given in a supplementary file. Section 4 provides an empirical illustration. A brief summary is given in Section 5.

2 German Electricity Price Densities and Dynamic Factors

We apply and investigate the multi-factor GMU method to a time series of hourly German electricity price data from 2012 to 2016. We chose the daily peak and trough at 19:00 and 03:00 plus an intermediate midday, to represent the range with three different types of hours. In this context, these are a viable representation of the price data for all the hours in a day. Since we wanted to do a direct comparison with previously published results, we did not use data beyond 2016. It is the methodological comparison that is the main objective of the research; even now many researchers are choosing not to use recent data since the beginning of 2020 because of the pandemic lockdowns, and their impact on the accuracy of the data. We set aside the final 10% of the data as a validation data set, while the remaining were used in the estimation phase.

The choice of German data was partly because it is one of the largest, most liquid and actively traded wholesale power markets in the world, and partly because it has attracted some of the most innovative price modeling research. One of the reasons it has attracted so much research on dynamic factors is because it has been prominent in the energy transition to renewable resources, particularly wind, and it has become evident (Gianfreda and Bunn (2018)) that when hourly wind production is high it causes hourly price levels to reduce, variance to increase and skewness to possibly flip from positive to negative. When the wind drops, the reverse effects, typically, happen to those moments. Thus, day-ahead wind forecasts are drivers of at least the first three moments in the day ahead hourly price density predictions.

For example, consider Figure 1, in which are plotted the German hourly price series in Euros/MWh for hour 12 during the years 2012 and 2016. Up until around 2012, the power system was dominated by fossil fuels (mainly coal and some gas) and the 2012 series exhibits periods of volatility clustering and positive spikes, typical of conventional power markets prior to the renewable transition. Positive spikes occur during periods of high demand and generation scarcity. In later years, the same power system operated with increasing wind and solar facilities replacing the gas and coal. The price evolution in the 2016 series shows that the underlying data density could have positive or negative skewness, depending on the hour of day. Negative spikes occur when there is too much wind generation, partly because of the subsidies received by the wind generators for generating and partly because inflexible generators would prefer to pay to generate rather than switch off and then on again in short succession. It is these dynamic effects that motivated this modeling research.

These time series, as with the time series in most of the published re-

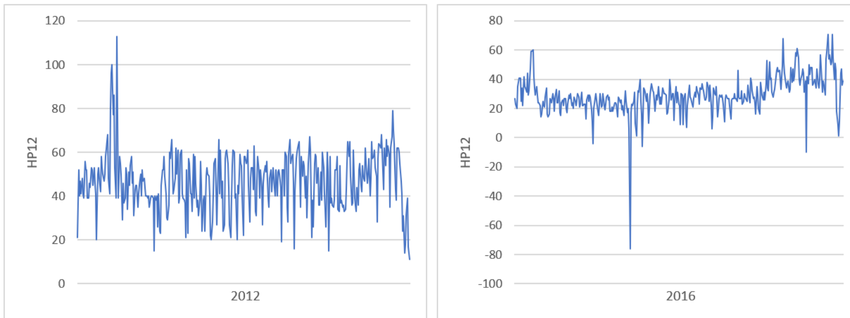

Figure 1: Daily time series for electricity prices (Euros/MWh) for hour 12 in 2012 and 2016. Similar plots can be obtained for hour 3 and hour 19. These are the time series data from 2012 to 2016 that are modeled in the paper, using a Bayesian GMU and skewed student-t family of densities.

search on daily wholesale electricity prices, consist of prices for delivery of power over hourly periods, established from competitive auctions at noon on the previous days. Generators make offers, and retailers make bids, for power for each hour of the following day, resulting in a set of 24 demand and supply functions which go into the day-ahead auction. Their intersections determine the market clearing prices and these are displayed daily on the power exchange website (EPEX SPOT, www.eex-group.com). From a modeling perspective, two aspects of this process are important. Although data are archived as a series of hourly prices, they are actually panels of 24 prices produced simultaneously at noon each day for the following day. So, the time series is actually a daily series of a 24-hour vector. This means that most modeling is done on a daily time series for specific hours (as in Figure 1). The markets are transparent in their fundamentals, and so all participants, prior to submitting their offers and bids each morning, will have access to data such as latest prices, fuel prices, demand data, weather forecasts, wind and solar generation forecasts, recent generator availability, transmission constraints, import/export of power, etc. This is a set of exogenous factors which all participants can reasonably be assumed to consider when formulating their individual bids and offers. The EU mandated the creation of a Transparency Platform (https://transparency.entsoe.eu/) to facilitate efficient trading across all member states and to limit the information advantages that large players held.

In general, power prices are well understood in terms of their fundamental and stochastic drivers. Power price formation is influenced by technological specificities, economic activity, social behavior and public policy. The underlying fuel commodities of gas and coal have been basic to production, whilst demand and supply are fundamental to price-clearing and operational constraints, as well as market conduct and regulatory interventions make significant impacts. Hourly, daily and seasonal periodicities are strongly evident in electricity demand, and these translate to some degree into prices. These factors are known, but the density forms specifying their interactions and dynamics remain active topics of research, especially with respect to robust forecasting and density prediction. For example, Lucia and Schwartz (2002), Knittel and Roberts (2005), Panagiotelis and Smith (2008), Chen and Bunn (2010), Aïd et al. (2013), Weron (2014) and Benth et al. (2015) show a range of fundamental formulations while De Vany and Walls (1999), Haldrup and Nielsen (2006), Koopman et al. (2007), Bunn and Gianfreda (2010), Escribano et al. (2011), Nan et al. (2014) advocate various predictive models. Price distributions can also vary by time-of-day; see, Damien et al. (2019) and Ekin et al. (2020). In terms of our focus, dynamic changes to skewness in particular are attracting attention, as they have obvious implications for tail risks. Outside the power applications, research on skewed distributions is vast and has found wide use in business, economics and finance applications. In econometrics, one of the earliest contributions was by Hinkley and Revankar (1977). Jones (2015) and Villa et al. (2019), as well as many references therein, discuss several families of skewed distributions and their applications.

Recognizing the relevance of complex features underlying energy price formations, Gianfreda and Bunn (2018) developed a family of stochastic latent moment models to better estimate the relationships of exogenous factors on the first four moments of the price data densities. They use their model to forecast the day-ahead price densities at various hours of the day. Following comprehensive comparisons, using five different types of skewed densities, EGARCH, and Quantile Regressions to model energy prices, they showed that using a multi-factor skewed student-t distribution for the electricity price data density resulted in superior overall performance. The specific skewed student-t representation on which they base their general multi-factor model is given by:

$$f_Y(y|\mu,\sigma,\nu,\tau) = (2/\sigma) f_{Z_1}(z) F_{Z_2}(\omega) \qquad y \in (-\infty,\infty) \tag{2}$$

where$-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$ and $\tau > 0$. And $z = (y-\mu)/\sigma$, $\omega = \nu\lambda^{1/2}z, \lambda = (\tau+1)/(\tau+z^2)$, f_{Z_1} is the pdf of a t-distribution with $\tau > 0$ degrees of freedom and F_{Z_2} is the cumulative distribution function of a t distribution with $\tau + 1$ degrees of freedom. The parameters μ, σ, ν, τ, relating to the first four moments, were linear functions of exogenous factors.

Following this, Abramova and Bunn (2020) undertook an extensive computational search of specifications to forecast the intra-day spread densities between prices at different hours, again on German data. This was relevant to the intra-day arbitrage trading of battery storage operators. They found that while the multi-factor skewed-t was most often best, the choice is not straightforward and other densities performed better at various hours.

Motivated by this, it is evidently desirable to move away from the particular density specification problem and to develop a formulation of the hourly price densities as realizations of a general underlying parametric stochastic process. The parameters of such a process should remain conditional on the same fundamental exogenous factors, and we achieve this with a stochastic, multi-factor moment model, using the GMU. We model the location (μ), scale (β) and skewness (λ) of the GMU (see equation (1)) via stylized regressions. In order to make a direct comparison with the three-moment, multi-factor skewed-t, which performed best in Gianfreda and Bunn (2018), we use a comparable set of factors in the linear regressions for each moment, and also focus on the same three daily time series for hours 3, 12 and 19. These hours were also selected, because they cover the entire spectrum of prices in a 24-hour period subject, of course, to some variability.

Let y_t denote the hourly electricity price for a particular hour on day t. Then, y_{t-1} is the autoregressive component which would impact y_t; most researchers have found that an adaptive lag of one day to be significant in power prices. The market clears with generator offers intersecting the retailer bids. The retailer bid functions will depend upon their demand forecasts. Moreover, the generators will make offers that are an increasing function of demand as more expensive units get called upon during peak demand periods. So prices will depend upon demand forecasts. Demand is usually inferred from electrical load measured on the system, and while load data is fully transparent, demand forecasts are not. In the absence of load forecasts, we follow a common practice in German price modeling of using the load at the same hour on the previous day as proxy for the retailers' forecast. Thus, the electricity wholesale load data (in thousands of MW) at time $t-1$, denoted $load_{t-1}$, would influence energy price formation at t. In contrast to load, the system operators provide forecasts for wind and solar generation (in thousands of MW), denoted $fwind_t$ and $fsolar_t$, respectively, and these are available to the market prior to the price auction at time t. Similarly, lagged prices of coal and gas, denoted $coal_{t-1}$, gas_{t-1}, would also impact prices at t. Finally, there is considerable practical evidence that points to varying price formations on German holidays and weekends. To incorporate this assumption, we add a dummy variable hol_t.

In order to get a sense of the flexibility of the density fits, Figure 2 shows both the GMU and skewed-t density fits for the hours 3,12,19 in 2012 and 2016. Visually, the GMU looks to fit better, but we need to investigate how the GMU conditional multi-factor model helps in better estimation, as specified in the next section; indeed, we show that the GMU fit is superior to the skewed-t, using pinball loss estimates of the extreme quantiles, as well as the root mean square error. We note at the outset that the primary focus is the Bayesian estimation of the parameters of the various regressions. These estimates would help participants quantify the impacts on the electricity price densities of the exogenous factors in the German market, and thereby evaluate price risk mitigations that could be achieved through hedging or forecasting those factors better.

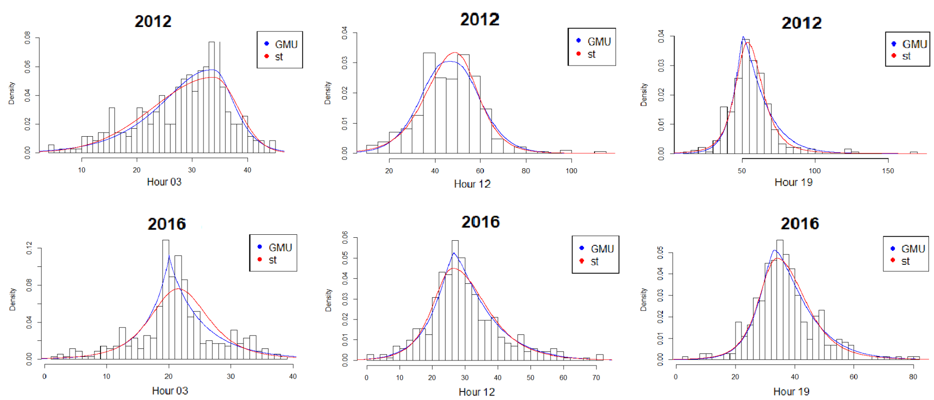

Figure 2: GMU and skewed student-t density fits for electricity prices (Euros/MWh) for hours 3,12,19 in 2012 (first row) and 2016 (second row). We also fit the skewed student-t model to our data, since the benchmark model of Gianfreda and Bunn (2018) uses this family of densities.

2.1 The GMU stochastic latent moment model

For the t^{th} daily observation, let

$$y_t \sim \text{GMU}(\mu_t, \alpha, \beta_t, \lambda_t). \tag{3}$$

The likelihood function for n observations with latent variable, $\mathbf{z}$, is given by

$$l(\mu_t, \alpha, \beta_t, \lambda_t; \mathbf{y}, \mathbf{z}) = \prod_{t=1}^{n} \frac{\beta_t^{\alpha} \exp(-\beta_t z_t)}{\{\exp(\lambda_t) + \exp(-\lambda_t)\}\Gamma(\alpha)} \left[z_t^{\alpha-2} \mathbb{1}\{z_t \geq r(y_t - \mu_t, \lambda_t)\} \right] \tag{4}$$

with

$$r(y_t - \mu_t, \lambda_t) = \max \Big\{ -(y_t - \mu_t) \exp(\lambda_t), (y_t - \mu_t) \exp(-\lambda_t) \Big\}.$$

To model the moments with dynamic factors, we denote $\mu_t, \lambda_t, 1/\beta_t$ in (4) with the following regressions, where the exogenous factors are the ones described above:

$$\begin{aligned} \mu_t = a_1 + \gamma_1 y_{t-1} + b_{11} hol_t + b_{12} load_{t-1} + b_{13} fwind_t + b_{14} fsolar_t \\ + b_{15} coal_{t-1} + b_{16} gas_{t-1} \end{aligned} \tag{5}$$

$$\beta_t = e^{-(a_2 + b_{21} hol_t + b_{22} load_{t-1} + b_{23} fwind_t + b_{24} fsolar_t + b_{25} coal_{t-1} + b_{26} gas_{t-1})} \tag{6}$$

$$\begin{aligned} \lambda_t = a_3 + b_{31} hol_t + b_{32} load_{t-1} + b_{33} fwind_t + b_{34} fsolar_t \\ + b_{35} coal_{t-1} + b_{36} gas_{t-1}. \end{aligned} \tag{7}$$

For the parameters in the above GMU model, the following proper priors that reflect diffuse beliefs were used. For all parameters in the mean regression, $(a_1, b_{11}, \gamma_1, \cdots b_{16})$, we used uniform prior distributions since the likelihood function involves only the indicator function. For all the other parameters (except a_2) we took a standard normal prior. Finally, simplifying terms in the above, we recognize that the density for a_2 is gamma distributed. Hence, a conjugate $gamma(1,1)$ prior was used.

The posterior joint distribution, clearly, is not analytically tractable. Hence, we use a Markov Chain Monte Carlo method, namely the Gibbs sampler (Gelfand and Smith (1990)) to provide full inference. Without loss of generality, a technical Appendix details the computational form of the multi-factor model; this is provided as a supplementary file. The algorithm was programmed in Python. The MCMC chain was run for $20,000$ iterations with a burn-in of $10,000$. Standard MCMC diagnostics were used to monitor convergence. In the interests of space, we only provide the Geweke diagnostics in the supplementary file.

3 Data Analysis

We fit our Bayesian multi-factor GMU and the benchmark multi-factor skewed student-t models to the data. The volume of output that the Bayesian GMU analysis of these data yields is quite vast since for every hour in a 24-hour cycle we can generate posterior distributions for all the parameters. This becomes unwieldy very quickly. Hence, we present critical insights for hours 3, 12 and 19 whose density estimate plots were shown in Figure 2. Our main aim is to compare our GMU results with the best skewed student-t model (equation (2)) from Gianfreda and Bunn (2018); these authors use maximum likelihood estimation for their skewed-t model, and so we do the same for our version of that model. (With non-informative priors, the Bayes and MLE estimates of the skewed-t model for these data are approximately the same.)

3.1 Parameter Estimation

The factor coefficient estimation results are shown in Table 1. In each cell, the left (right) sign corresponds to the GMU (skewed student-t) model. This table contains the high-level summary of all the coefficients from the GMU and skewed student-t latent moments model. (The corresponding full posterior summaries for the GMU model are available on request.) The results are consistent with expectations for most part. Note that y_{t-1} only appears in the mean moment equation. Likewise, solar is zero in the morning hours; hence there are no entries corresponding to this variable for Hour 3 in the table.

	Hour 3			Hour 12			Hour 19		
	μ	β	λ	μ	β	λ	μ	β	λ
hol_t	- −	+ +	- −	- −	- +	- −	- −	- −	- −
y_{t-1}	+ +			+ +			+ +		
$load_{t-1}$	+ +	- −	+ +	- +	- −	- −	- +	- +	+ +
$fwind_t$	- −	+ +	- −	- −	- +	- −	- −	+ −	- −
$fsolar_t$				- −	- +	- −	- −	- −	- −
gas_{t-1}	+ +	+ +	- −	+ +	+ +	- +	+ +	+ +	- −
$coal_{t-1}$	- −	+ +	+ +	+ +	- +	+ +	+ +	- −	+ +

Table 1: High level summary of signs for mean (μ), scale (β), and skewness (λ) regressions: in each cell left (right) sign is from the GMU (Skewed student-t) model. The GMU signs for the coefficients are from the Bayesian algorithm developed in this paper, while the skewed student-t signs for the coefficients were estimated using maximum likelihood as in Gianfreda and Bunn (2018). We used our dataset from 2012-2016 to estimate the GMU and skewed student-t models. Also, these estimates were based on 90% of the data, since the last 10% were set aside for validation purposes.

The main result is that increasing both wind and solar production always reduced the level and skewness of hourly prices across all three hours, as expected. Of course, solar does not impact price in hour 3. We also found, as expected, that price levels and skewness decline during holidays/weekends. The lagged price coefficients in the mean equations are intuitively positive given the known adaptive element in power market bids and offers. Similarly the lagged load coefficients in the mean equations are positive, consistent with increased demand forecasts driving up price levels. Intuition for higher moment effects is not obvious and there may be interactions. Interactions between the fuel prices, gas and coal, are well known from supply function dynamics: coal and gas are direct substitutes and often switch in their power market price setting roles according to their relative commodity prices. During the period 2012-2016, gas prices fluctuated whilst coal prices declined. Higher gas prices increased electricity prices and volatility, as expected. The coal effects are less straightforward. Thus, higher coal increased electricity prices during the day, but reduced them at night. This may seem counter-intuitive, the explanation being that higher coal prices put coal generation on the margin, and as the marginal generators during the low demand period (Hour 3), being inflexible, they will reduce their prices in order to avoid being called off. Without going into further specific interpretations, however, the signs are generally plausible and consistent across the two methods. Our main focus is upon the clear effect of wind and solar forecasts on the day-ahead price skewness.

It is perhaps less relevant to compare the coefficients to those in Gianfreda and Bunn (2018), although again they are broadly consistent. However, they used data from 2007 to 2011, whereas our time horizon is more recently relevant, being 2012 through 2016. They estimated the kurtosis of the various densities in their model but concluded that "intuitions regarding their [parameter] signs are equivocal, and the gains in fit are small." We agree since we also modeled kurtosis but found that the gains were non-existent in terms of predictability. This should not come as a surprise, since once the volatility and skew are modeled, they impact the peakedness of the underlying data density as well. Since most of the discrepancies in the coefficient signs between the two studies related to variance, the use of kurtosis in their study would be an explanation. They also used CO_2 as a covariate and found that it was not as meaningful. Hence, we omitted it in our model. Lastly, in the various comparisons they executed, they also fitted a VAR component to the latent moments but they concluded that these models may result in over-fitting, despite their conceptual appeal. Nevertheless, there is broad consistency between that study and ours, particularly on the skewness effect. However, they do not model risk estimation via probability forecasts and merit-order effects which we do. Thus, the above table should be seen as a first step toward the overall goal of risk estimation from a Bayesian perspective.

Out-of-sample Predictability As a quality test, we bench-marked predictions from the GMU against the multi-factor skewed student-t model. As noted earlier in the paper, we set aside the last 10% of the data for each of the three hours discussed in this paper. Consider Table 2. It shows the root mean square error (RMSE) for the validation data set, and the pinball loss estimates at the 95th and 99th percentiles for Hours 3, 12 and 19. It is clear that both from a point-forecast (RMSE) and a density quantile forecast the Bayesian GMU outperforms the skewed student-t.

Convergence Tests. For the Bayesian GMU model, we monitored the convergence of the posterior distributions of the regression parameters using trace plots and Geweke convergence diagnostics. In the interests of space, we provide the latter in the supplementary file.

Hour	RMSE	Pinball Loss at the 95th and 99th Percentiles	
	GMU (Skewed-t)	GMU (Skewed-t)	GMU (Skewed-t)
3	6.92 (8.12)	5.03 (6.28)	5.23 (6.52)
12	9.89 (10.58)	4.13 (4.20)	4.22 (4.30)
19	12.63 (16.03)	7.84 (11.27)	8.14 (11.73)

Table 2: RMSE and Pinball Loss Comparison

3.2 Posterior distributions for the Merit-Order Effect

One of the advantages of the Bayesian approach in this application is that it provides a complete quantification of the “Merit-order Effect (MOE)” in electricity price formation. The MOE is the process by which fluctuations in the available capacity of “inframarginal” generation translate into higher price volatility. Since wholesale electricity prices are set by the market clearing technologies, which have higher short-run marginal costs than the more efficient (inframarginal) generators, changes in the inframarginal outputs will cause the market clearing prices to change by much than their own inframarginal costs. Thus, fluctuations in the output of renewable facilities such as wind, which have zero marginal cost (or negative if subsidies are included) could translate to much higher price volatility if gas generation is the marginal technology. From a statistical perspective, we are able to fully quantify the MOE using the posterior distributions of the appropriate parameters from our multi-factor GMU model. For the mean effects, Equation 5 provides these parameters which are essential inputs into many risk management and hedging models. Thus, the commonly used "delta" hedging needs to know how price volatility may be reduced if the uncertainty in one of the driving factors can be mitigated. For example, gas prices can be hedged through financial options, or wind outputs through weather derivatives. Equation 5 gives the effects on mean prices of such hedging and thereby helps to value its risk mitigation. If such hedges are not available, or too expensive, the coefficients from Equation 5 will at least indicate the value of better forecasts of these factors.

The benefit of the Bayesian estimation is that we have posterior distributions for the factor parameters and this will allow more precise risk mitigation valuations. Consider Figures 4, 5 and 6. These summarize with boxplots the posterior distributions of the mean regression parameters for Hours 3, 12 and 19, respectively. The parameters correspond to the variables wind (b_{13}), solar (b_{14}), coal (b_{15}) and gas (b_{16}), respectively. Looking at Figure 5, for midday, we see the negative coefficients for wind and solar output forecasts, as we saw in Table 1 and would expect from increased output. But the new insight here is the slightly more negative and much less dispersed effect of solar compared to wind. It is not obvious why the marginal effect of wind output on prices should have a more dispersed posterior distribution than solar, since they are both zero marginal cost, inframarginal technologies. The answer will likely be in the market specificities. In the German case it will be due, at least in part, to the fact that wind resources are mostly in the north and solar in the south, with transmission constraints particularly on north to south transfers. Constraints will lead to higher prices in a nonlinear way, and so we see the risk of a much more negative coefficient for wind compared to solar. An implication of this is that it might be more effective to hedge (or improve forecasts for) solar compared to wind. Gas is hedgeable with financial options and in Figure 6 we see the strong effect of gas prices on the evening peak, since gas turbines generally provide the peaking technology. The posterior distribution of the mean has some negative skewness and will therefore influence the hedging formulas, distinct from the usual normality assumption.

Non-normality in the price risk itself, rather than in the mean, is however more thoroughly analyzed through inspecting the higher moments, especially the skewness via Equation 7. Figures 7 through 9 show, for example, the posterior box plots for the coefficients of the skewness regression appearing in Equation 7. Recall that in looking at midday in Figure 5, we observed the slightly more negative and much greater dispersed mean coefficient effect of wind compared to solar, whilst in Figure 8, again for midday, we have extra insight into the greater dispersion of this wind coefficient. We see in Figure 8 that the skewness coefficient is larger in absolute value for wind than for solar. In other words wind fluctuations not only cause greater price fluctuations than solar, but also a longer negative tail (higher risk of very low prices). This is a manifestation of the transmission congestion in northern Germany, where the wind resources are located, creating local excess supply which can occasionally lead to extreme price cuts by inflexible generators that need to stay running for technical reasons. This type of insight is one of the merits of the Bayesian approach to better model the MOE, and has practical value in quantifying the benefits of hedging the exogenous factors (e.g. through wind derivatives).

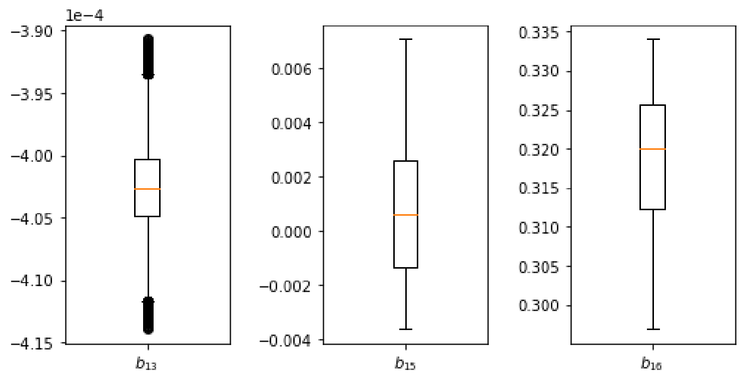

Figure 3: Hour 3 Boxplots for b_{13}, b_{15}, b_{16} from Equation 5

$\mu_{b_{13}}$ = -0.0004 $\quad \mu_{b_{15}} = 0.0008 \quad \mu_{b_{16}} = 0.3188$

$sd_{b_{13}} = 3.4257e^{-6} \quad sd_{b_{15}} = 0.0025 \quad sd_{b_{16}} = 0.0082$

$iqr_{b_{13}} = 4.5325e^{-6} \quad iqr_{b_{15}} = 0.0039 \quad iqr_{b_{16}} = 0.01347$

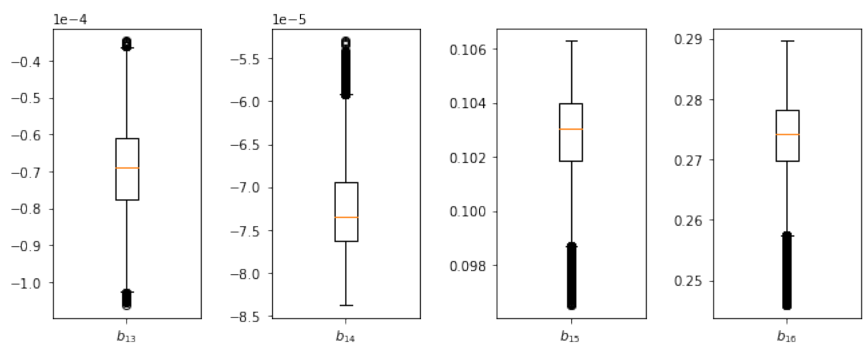

Figure 4: Hour 12 Boxplots for $b_{13}, b_{14}, b_{15}, b_{16}$ from Equation 5

$\mu_{b_{13}} = -6.9323e^{-5} \quad \mu_{b_{14}} = -7.2780e^{-5} \quad \mu_{b_{15}} = 0.1028 \quad \mu_{b_{16}} = 0.2738$

$sd_{b_{13}} = 1.1790e^{-5} \quad sd_{b_{14}} = 4.8085e^{-6} \quad sd_{b_{15}} = 0.0017 \quad sd_{b_{16}} = 0.0070$

$iqr_{b_{13}} = 1.6572e^{-5} \quad iqr_{b_{14}} = 6.8347e^{-6} \quad iqr_{b_{15}} = 0.0021 \quad iqr_{b_{16}} = 0.0083$

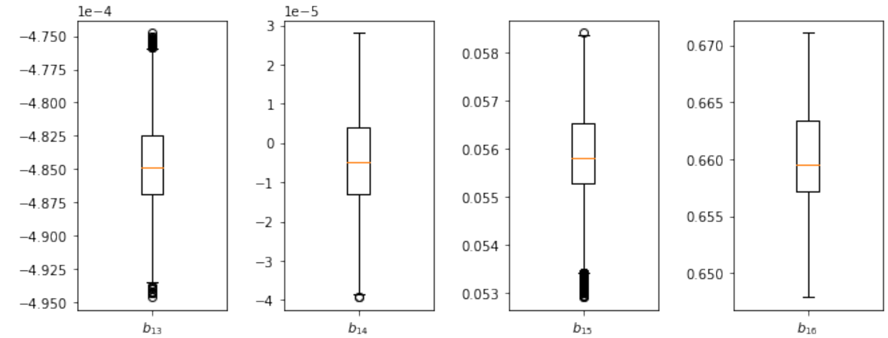

Figure 5: Hour 19 Boxplots for $b_{13}, b_{14}, b_{15}, b_{16}$ from Equation 5

$$\mu_{b_{13}} = -0.0004 \quad \mu_{b_{14}} = -4.4298e^{-6} \quad \mu_{b_{15}} = 0.0559 \quad \mu_{b_{16}} = 0.6601$$

$$sd_{b_{13}} = 3.1795e^{-6} \quad sd_{b_{14}} = 1.1978e^{-5} \quad sd_{b_{15}} = 0.0010 \quad sd_{b_{16}} = 0.0044$$

$$iqr_{b_{13}} = 4.4156e^{-6} \quad iqr_{b_{14}} = 1.7251e^{-5} \quad iqr_{b_{15}} = 0.0012 \quad iqr_{b_{16}} = 0.0063$$

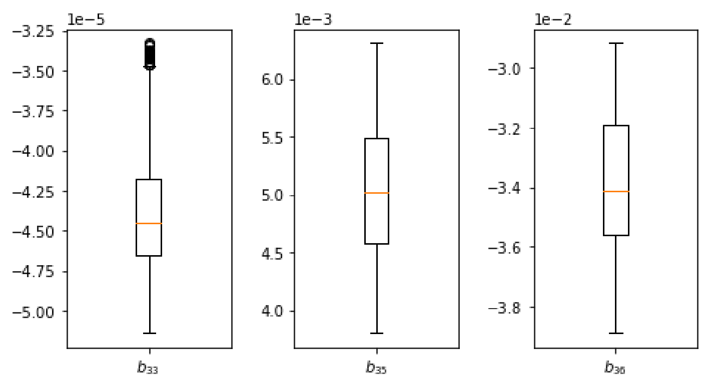

Figure 6: Hour 3 Boxplots for b_{33}, b_{35}, b_{36} from Equation 7

$$\mu_{b_{33}} = -4.412e^{-6} \quad \mu_{b_{35}} = 0.005 \quad \mu_{b_{36}} = \text{-}0.0338$$

$$sd_{b_{33}} = 3.145e^{-6} \quad sd_{b_{35}} = 0.0005 \quad sd_{b_{36}} = 0.0022$$

$$iqr_{b_{33}} = 4.745e^{-6} \quad iqr_{b_{35}} = 0.0009 \quad iqr_{b_{36}} = 0.00368$$

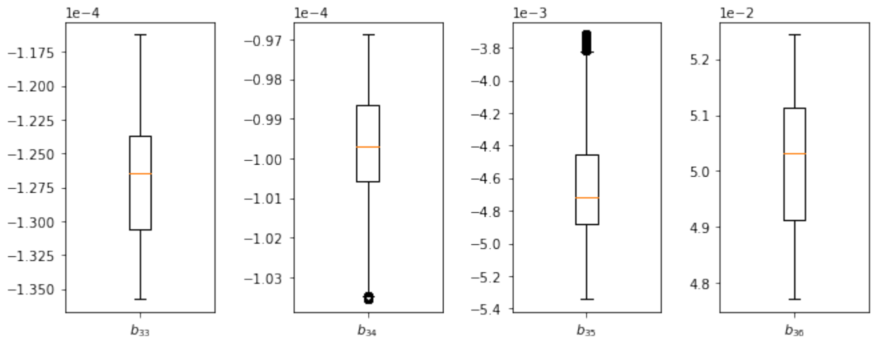

Figure 7: Hour 12 Boxplots for $b_{33}, b_{34}, b_{35}, b_{36}$ from Equation 7

$\mu_{b_{33}} = -0.0001$ $\mu_{b_{34}} = -9.971e^{-5}$ $\mu_{b_{35}} =$ -0.005 $\mu_{b_{36}} = 0.0501$

$sd_{b_{33}} = 4.756e^{-6}$ $sd_{b_{34}} = 1.375e^{-6}$ $sd_{b_{35}} = 0.0004$ $sd_{b_{36}} = 0.0011$

$iqr_{b_{33}} = 6.941e^{-6}$ $iqr_{b_{34}} = 1.92e^{-6}$ $iqr_{b_{35}} = 0.0004$ $iqr_{b_{36}} = 0.002$

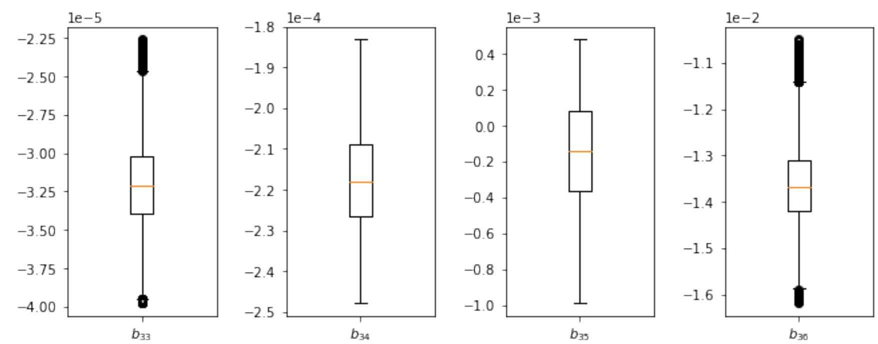

Figure 8: Hour 19 Boxplots for $b_{33}, b_{34}, b_{35}, b_{36}$ from Equation 7

$\mu_{b_{33}} = -3.179e^{-5}$ $\mu_{b_{34}} = -0.0002$ $\mu_{b_{35}} =$ -0.00015 $\mu_{b_{36}} =$ -0.0135

$sd_{b_{33}} = 3.354e^{-6}$ $sd_{b_{34}} = 1.206e^{-5}$ $sd_{b_{35}} = 0.00031$ $sd_{b_{36}} = 0.00107$

$iqr_{b_{33}} = 3.709e^{-6}$ $iqr_{b_{34}} = 1.742e^{-5}$ $iqr_{b_{35}} = 0.00045$ $iqr_{b_{36}} = 0.00111$

4 Summary and Conclusions

This research has been motivated by the recognition that the predictions of higher order moments, skewness in particular, is finding increasing applications in financial engineering and risk analysis generally. Electricity prices are an important example whereby the data densities of day-ahead prices can flip between positive and negative skewness according to the weather conditions for renewable power generation. The general problem with estimating latent stochastic moments, especially with exogenous factors, has been the selection of appropriate density functions to represent the data. That has been an issue for many years in GARCH-X modeling for volatility and it is more recently emerging as an awkward consideration for stochastic skewness. Recently, Kang et al. (2020) developed a novel gamma mixture of uniform (GMU) densities that outperform various families of densities including skewed student-t, skewed-normal, asymmetric Laplace, etc. We advance the GMU family by estimating its higher stochastic moments, particularly skewness, in which the moments are conditionally dependent upon several exogenous factors. We constructed a Bayesian version of the model and used a Markov Chain Monte Carlo scheme for its implementation. Empirical testing showed that it outperforms the multi-factor skewed student-t family of densities, previously advocated in this context. Additionally, using Bayesian estimation the new methodology provides a complete description of the uncertainty in the estimation of the coefficients for those exogenous factors. Empirical testing on day-ahead hourly electricity prices in the German market, where renewable energy sources, such as wind and solar, play a critical role in the formation of electricity price risk, validated the extra predictive accuracy of this formulation, using the pinball loss function. The posterior distributions of the exogenous factor parameters show non-normality and can therefore provide more precise inputs to hedging models for electricity price risk management. Indeed, using the Bayesian output, it would be interesting to develop stylized stochastic optimization algorithms to find optimal hedging strategies. We leave this for future research. In conclusion, we believe the GMU methodology to model stochastic skewness could prove useful in many potential applications in financial and commodity risk estimations.

References

Aboura, S. and Maillard, D. (2016). Option pricing under skewness and kurtosis using a Cornish–Fisher expansion. *Journal of Futures Markets*, 36(12):1194–1209.

Abramova, E. and Bunn, D. (2020). Forecasting the intra-day spread densities of electricity prices. *Energies*, 13(3):687.

Aïd, R., Campi, L., and Langrené, N. (2013). A structural risk-neutral model for pricing and hedging power derivatives. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 23(3):387–438.

Benth, F. E., Christensen, T. S., and Rohde, V. (2020). Multivariate continuous-time modeling of wind indexes and hedging of wind risk. *Quantitative Finance*, pages 165–183.

Benth, F. E., Lange, N., and Myklebust, T. A. (2015). Pricing and hedging quanto options in energy markets. *Journal of Energy Markets*, 8(1).

Borland, L. and Bouchaud, J.-P. (2004). A non-Gaussian option pricing model with skew. *Quantitative Finance*, 4(5):499–514.

Brusaferri, A., Matteucci, M., Portolani, P., and Vitali, A. (2019). Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices. *Applied Energy*, 250(C):1158–1175.

Bunn, D. W. and Gianfreda, A. (2010). Integration and shock transmissions across European electricity forward markets. *Energy Economics*, 32(2):278–291.

Canelas, E., Pinto-Varela, T., and Sawik, B. (2020). Electricity portfolio optimization for large consumers: Iberian electricity market case study. *Energies*, 13(9).

Chen, D. and Bunn, D. W. (2010). Analysis of the nonlinear response of electricity prices to fundamental and strategic factors. *IEEE Transactions on Power Systems*, 25(2):595–606.

Damien, P., Fuentes-García, R., Mena, R., and Zarnikau, J. (2019). Impacts of day-ahead versus real-time market prices on wholesale electricity demand in Texas. *Energy Economics*, 81:259–272.

de Jong, C. and Huisman, R. (2000). From skews to a skewed-t: Modelling option-implied returns by a skewed student-t. In *Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr)(Cat. No. 00TH8520)*, pages 132–142. IEEE.

De Vany, A. S. and Walls, W. D. (1999). Cointegration analysis of spot electricity prices: insights on transmission efficiency in the western US. *Energy Economics*, 21(5):435–448.

Ekin, T., Damien, P., and Zarnikau, J. (2020). Estimating marginal effects of key factors that influence wholesale electricity demand and price distributions in Texas via quantile variable selection methods. *Journal of Energy Markets*, (Forthcoming).

Escribano, A., Ignacio Peña, J., and Villaplana, P. (2011). Modelling electricity prices: International evidence. *Oxford Bulletin of Economics and Statistics*, 73(5):622–650.

Gelfand, A. and Smith, F. A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.

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

Giesecke, K., Kim, B., Kim, J., and Tsoukalas, G. (2014). Optimal credit swap portfolios. *Management Science*, 60(9):2291–2307.

Hagfors, L. I., Kamperud, H. H., Paraschiv, F., Prokopczuk, M., Sator, A., and Westgaard, S. (2016). Prediction of extreme price occurrences in the German day-ahead electricity market. *Quantitative Finance*, pages 1929–1948.

Haldrup, N. and Nielsen, M. Ø. (2006). A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1-2):349–376.

Hinkley, D. and Revankar, N. (1977). Estimation of the Pareto law from underreported data: a further analysis. *Journal of Econometrics*, 5(1):1–11.

Jondeau, E. and Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management*, 12(1):29–55.

Jones, M. (2015). On families of distributions with shape parameters (with discussion). *International Statistical Review*, 83(2):175–192.

Kang, L., Damien, P., and Walker, S. (2020). On a transform for modeling skewness. *Brazilian Journal of Probability and Statistics. To Appear.*

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

Koopman, S. J., Ooms, M., and Carnero, M. A. (2007). Periodic seasonal reg-arfima–garch models for daily electricity spot prices. *Journal of the American Statistical Association*, 102(477):16–27.

Lucia, J. J. and Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the Nordic power exchange. *Review of Derivatives Research*, 5(1):5–50.

Nan, F., Bordignon, S., Bunn, D. W., and Lisi, F. (2014). The forecasting accuracy of electricity price formation models. *International Journal of Energy and Statistics*, 2(01):1–26.

Panagiotelis, A. and Smith, M. (2008). Bayesian density forecasting of intraday electricity prices using multivariate skew-t distributions. *International Journal of Forecasting*, 24(4):710–727.

Serinaldi, F. (2011). Distributional modeling and short-term forecasting of electricity prices by generalized additive models for location, scale and shape. *Energy Economics*, 33(6):1216–1226.

Villa, C., Leisen, F., and Rossini, L. (2019). Loss-based approach to two-piece location–scale distributions with applications to dependent data. *Statistical Methods and Applications*, Open access:DOI: https://doi.org/10.1007/s10260–019–00481–x.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081.