Unveiling endogeneity and temporal dependence in energy prices and demand in Iberian countries: a stochastic hidden Markov model approach

Jorge Antunes¹ · Luis Alberiko Gil-Alana² · Rossana Riccardi³ · Yong Tan⁴ · Peter Wanke¹

Accepted: 19 July 2021 / Published online: 29 July 2021
© The Author(s) 2021

Abstract

In this paper, we analyze the temporal dependence in energy prices and demand using daily data of Portugal and Spain over the period 2007–2017. The methodology used is based on a stochastic Hidden Markov Model and the results indicate first that all significant relationships between energy prices and demands were found to be positive; second, spot prices are only time dependent on future prices and spot energy, while future energy is solely time dependent on spot energy behavior; third, future prices are not only autocorrelated but also time-dependent with spot energy and future energy demands level; and finally, spot energy is autocorrelated and time-dependent with future prices and future energy. Policy implications of the results obtained are presented at the end of the article.

Keywords Renewables · Decarbonization · Iberian countries · Energy demand · Energy prices · Stochastic HMM · Copulas · Bootstrapped VAR models

JEL Classification C22

1 Introduction

Energy is an important resource for individuals, business entities as well as the macro-economic development. Individuals use energies such as electricity, gas and petrol on a daily basis, while business entities reply to a large extent on different energy resources

✉ Yong Tan
a.y.tan@hud.ac.uk

¹ COPPEAD Graduate Business School, Federal University of Rio de Janeiro, Rua Paschoal Lemme, 355, Rio de Janeiro, Brazil

² Faculty of Economics, University of Navarra, Pamplona, Spain

³ Dipartimento di Economia e Management, Università Degli Studi di Brescia, Via S. Faustino 74/b, 25122 Brescia, Italy

⁴ Department of Accounting, Finance and Economics, Huddersfield Business School, University of Huddersfield, Huddersfield HD1 3DH, Queensgate, UK

for their daily operations; finally, the macroeconomic development fueled by the transportation and logistics industries cannot be operated without energy. The price of the energy, obviously, is a factor that is concerned by different parties of the economy and is an important determinant of individual activities, industry-related activities as well as economic activities as a whole.

The spot price in the energy sector mainly means the price that retailers pay to the wholesales which experiences a level of volatility based on the demand and supply (Kaleta and Toczyłowski, 2008), whereas futures contract is used by the investors to buy or sell energy at a price and date which are predetermined to reduce the price fluctuations (Huisman et al., 2009). Over the past few decades, there are a number of studies investigating the spot and futures prices in the energy sectors, (see literature review below in detail), very few pieces of research considered the spot demand and futures demand at the same time.

In the first half of the twentieth century, the Iberian energy sector, namely the energy sector of Portugal and Spain, was different from the one of other European or OECD countries from the perspective of its higher degree of dependence on external energy. Now Spain and Portugal are the two countries leading the energy transition promoted by electricity generation from renewable energy sources (Perez-Franco et al., 2020). There are several electricity markets among the European countries, one of which is operated by OMIE for the Iberian electricity market, the main responsibility of which is to manage the day-ahead and intraday electricity markets in Spain and Portugal (Canelas et al., 2020). Because of the 2007–2009 financial crisis hit the economy of both Portugal and Spain very hard and the disastrous effect was even stronger than other European countries, therefore, both of these two countries have been attempting to make transition toward a more sustainable and a resilient economy. One of the efforts made by the governments was to reduce the level of pollution from energy production. In Particular, Portugal committed to have electricity generation from renewable resource accounting for 60% of total production by 2020 (Gil-Alana et al., 2020). In addition, investment in renewable energy will play a positive impact on stabilizing climate, increase employment and economic growth (Pollin et al., 2015), it will also reduce GDP per capita emission as well as diminish the energy dependency (Perez-Franco et al., 2020), which is of particular importance for the Iberian Peninsula. Through the investigation of the relationship between energy price and energy demand under the propose innovative method, we would be able to not only establish a reasonable price mechanism in the spot and futures energy market in Iberian market, but also the price mechanism will provide a guidance and motivation for these two countries to engage in the endeavor in the process of decarbonization in the energy sector by producing energy from coal sources towards sustainable ones such as wind and solar.

There are quite a few questions need to be answered which have not been addressed by the previous literature: (1) will energy spot price be related to the futures price, spot demand and futures energy demand? (2) Will there be a linkage between futures energy price and spot demand, spot energy price as well as the futures energy demand? (3) Will both futures and spot energy prices as well as futures energy demand influence the spot energy demand? (4) Is there a potential influence of spot and futures energy price as well as spot energy demand on futures energy demand?. Answers these questions would be very important due to the consideration that this will fill in the gap of the empirical literature on this topic, and in addition, the answers to these questions will provide accurate and clear information and guidance related to the pricing mechanism in the spot and futures energy market in Iberian Peninsula. This research idea can be applied or expanded to the whole European Market or other regional markets.

A Hidden Markov Model (HMM) is a type of graphical model often used to model temporal data. Unlike traditional Markov Models, HMMs assume that the data observed is not the actual state of the model, but is instead generated by underlying hidden states (Danisman & Kocer, 2021). HMMs have been widely applied to many different fields because of their flexibility and computational efficiency, such as energy (Cai et al., 2018; Ullah et al., 2018). The current study distinguishes itself from the previous literature in the energy sector by investigating the endogenous and temporal dependent relationship between energy prices and demands through a stochastic HMM derived from copula generation. The stochastic HMM method benefits from the advantages of being able to examine the temporal dependence and endogeneity on price and demand from the perspective of the initialization of state and probabilities of transition. Within this specific model, both Vector Auto Regressive (VAR) and Multinomial Logit (ML) models can be used under the bootstrapped procedure.

The Energy Information Administration of US department of Energy has developed a method called levelized avoid cost of electricity (LACE) to compare and select energy investment, this method compares the levelized costs (LCOE) with the avoided costs when implementing new projects. The net value (the difference between LACE and LCOE) can be thought of as the potential profit (or loss) per unit of energy production for the plant (EIA, 2013). Our proposed HMM model can be beneficial for a better predictability of the financial equations for assessing decarbonized alternative sources of energy, this being one of the motivations for the present work. This study departs from Gil-Alana et al. (2020) by digging further into the issue of endogeneity between prices and consumption in Iberian energy markets. While this previous paper used fractional integration to examine the degree of persistence of these series in spot and futures markets in Spain and Portugal, this research now uses a stochastic HMM approach to unveil the cause-effect relationships and feedback process that may exist within the temporal dependent structure previously analysed. By answering this remaining gap, it is possible to ascertain the better sequence of actions for improving market regulation and development in Iberian Peninsula. In addition, the objectives of this work are the following:

- (1) We investigate the autocorrelation of futures energy price and demand as well as the relationship between futures price and spot price. The investigation of this is very important due to the fact that the futures contracts play a role of price discovery and they are used to reduce particular risks (Nicolau & Palomba, 2015).
- (2) We examine the auto-correlation of spot energy price and energy demand. The examination of this issue is motivated by the fact that although the spot market only accounted for a very small percentage in the energy transactions, it has a degree of influence on the formation of government selling prices or official selling prices (Silvério & Szklo, 2012).
- (3) We evaluation the inter-dependence between energy price and energy demand in both the spot and futures market. The evaluation of this issue is very essential considering the fact that based on the energy demand, relevant policies can be made related to the price mechanism, while the price mechanism can be used as a regulatory tool to adjust the demand in the energy spot and futures market, which to a certain degree, contribute to the achievement of macroeconomic goals.

This paper will have the following five parts. Section 2 reviews the literature related to the spot and futures energy market, Sect. 3 presents the innovative methods adopted

in the current study. Results are analyzed and discussed in Sect. 4. Conclusions follow in Sect. 5. The R code developed for this research is disclosed in the Appendix for the sake of research reproducibility.

2 Literature review

Academic researchers have been active in research in the energy sector from different perspectives, one stream of the studies investigated energy efficiency and its determinants (Bashir et al., 2020; Bian et al., 2017; Lv et al., 2015; Omrani et al., 2019; Zhou et al., 2017), the second stream of studies examined the issue of energy demand and energy consumption (Li et al., 2017; Shadzahl, 2020; Shahzad et al., 2021; Talbi et al., 2020; Yang & Yang, 2015); while there is a growing awareness and concern during recent decades about climate change derived from carbon emission, therefore, relevant efforts have been made by academic research in the area of renewable energy and carbon emissions (Dogan et al., 2021; Farooq et al., 2019; Fatima et al., 2021; Ghazouani et al., 2020; Sarwar et al., 2019; Shahzad et al., 2020). Although relevant studies during recent couple of years focused on the evaluation of energy demand as illustrated before, however, none of the above-mentioned studies provided a detailed analysis about the energy spot market and futures market, let alone the inter-relationships among the price and demand in these two markets. However, over the past few decades, the literature has made effect in estimating the spot and futures energy markets from different perspectives. King and Cuc (1996) investigate the price convergence of the North American natural gas spot markets under the time-varying parameter analysis, which benefits from the ability to assess the strength of price convergence across different gas-producing basins, the other merit of the method lies to its explicit presence of time varying parameters. The findings show that the degree of price convergence becomes significantly higher after the price deregulation took place since the mid 1980s. Instead examining the degree of price convergence, Worthington et al. (2005) evaluate the transmission of spot electricity prices and price volatility in Australian under the multivariate generalised autoregressive conditional heteroskedasticity model. The findings suggest that the price volatility in some markets is affected by the shocks from other markets. Employing a VAR model, which has the advantage of allowing regularities in the data without imposing as many prior restrictions compared to structural model, Park et al. (2006) examine the relationship among 11 US spot market electricity prices. The findings show that the transmission effect is limited across different market in a long-time frame. The evidence of electricity market integration between Spain, Portugal, Austria, Germany, Switzerland and France is assessed by Ciarreta and Zarraga (2015). The relationship between spot prices in the national Australian market is investigated by Yan and Trück (2020) under the dynamic network analysis. the results from the principal component analysis as well as the granger causality network analysis show that there are few factors that are related to the derived measure of interdependence including unexpected high demand for electricity and sudden increase in the level of price volatility, among others.

The integration is mainly measured by the price convergence and spillovers. The investigation is facilitated by using multivariate Generalized Autoregressive Conditional Heteroscedastic models. They find dynamic correlations between Spain–Portugal, Germany–Austria and Switzerland–Austria. On the other hand, the integration between Spain–France and Germany–France is very weak.

Ewing et al. (2006) use three different markets in the United States (crude oil, heating oil and gasoline) to investigate the asymmetric relationship between spot price and futures price under the momentum-threshold autoregressive (M-TAR) model, the result suggests that the spot price and futures price for each of these three markets are cointegrated. The research of this topic was also undertaken by other studies. Maslyuk and Smyth (2009) use US WTI and the UK Brent to examine whether the crude oil spot and futures prices are cointegrated under a residual-based cointegration test. This method possesses the property of allowing for one structural break in the cointegrating vector and high frequency data. The findings show that the spot and futures crude prices under both the same grade and different grades are cointegrated.

Weron and Zator (2014) address the pitfalls of using the linear regression models to investigate the relationship between spot and futures prices in the electricity market (i.e. simultaneity problems, measurement errors and seasonality issues) by employing regression models with GARCH residuals, the results from the sample of a 13-year long spot and futures prices series in the Nord Pool electricity market suggest that there is a significant and positive influence of water reservoir level on the risk premium. Using the Dow Jones US commodity index, Beckmann et al. (2014) analyze the spot and futures prices in the energy sector. The long-run relationship between the spot and futures prices is examined in the first step followed by the examination of the adjustment pattern under a smooth transition model, which provide a higher degree of flexibility compared to the threshold model. The findings show that the present price discovery function is significantly influenced by the past relative volatility, while Zhang and Wang (2013) find that crude oil and gasoline futures market contribute more to the price discovery function compared to the spot energy market. This finding is also supported by Silvrio and Szklo (2012). The relationship between spot prices and futures prices of crude oils from North America, Europe, Africa and the middle East is investigated by Kaufmann and Ullman (2009) under two methods, namely a two-step DOLS error correction model and a vector error correction model. The findings show that the linkage between spot and futures markets is relatively weak. Nicolau and Palomba (2015) investigate the relationship between spot and futures prices of natural gas and oil using daily data from 7th January 1997 to 31st May 2014. The results from the recursive bivariate VAR models suggest that there is a level of interactions between spot and futures prices. More specifically, it is found that the gas price in Europe is significantly affected by the crude oil price (Asche et al., 2018).

The US natural gas spot and futures prices and their responses to storage change surprises are examined by Chiou-Wei et al. (2014). The results suggest that there is a negative relationship between changes in futures price and surprises in the change in the natural gas in storage. The study further indicates that the change in the price level will be firstly reflected in the futures market and then the information will be flown to the spot market.

Efforts have been made in the literature to investigate the predictability of natural gas spot and futures prices, in particular, using daily data of Henry Hub natural gas spot and futures prices, Mishra and Smyth (2016) answer the question of whether futures natural gas prices can predict the natural gas spot prices. The finding show that the predictability of futures natural gas prices is not better than the one predicted from the random walk model.

Therefore, as far as we are concerned, we are the pioneer to use multivariate copulas together with HMM to capture the joint variations of energy prices and demands in future and spot markets in a stochastic manner. Not only the multivariate copulas modelling adopted here can represent the times series being examined in a more precise manner, but they also allow follow-on robustness analysis by means of bootstrapping in terms of the

models (VAR and ML) used to unveil their time dependence and endogeneity. In summary, we have the following research hypotheses for the Iberian energy industry:

H1a Spot energy prices are autocorrelated.

H1b Spot energy prices are endogenously time-dependent with future energy prices, spot energy demand, and future energy demand.

H1a has been documented in the literature, as we can see that Park et al. (2006) present the same findings in terms of the relationship between spot prices in the US electricity market, therefore, we hypothesize that there would be the same case for the Iberian energy market. In terms of H1b, the causality runs from futures prices to spot prices is in accordance with Moose and Al-Loughani (1995), while Pindyck (2011) argues that most of the commodities including the energy, the price level will have a positive correlation with the condition of the macroeconomic environment due to the fact that an economic growth that will push up the demand level and further increase the level of energy price.

H2a Future energy prices are autocorrelated.

H2b Future energy prices are endogenously time-dependent with spot energy prices, spot energy demand, and future energy demand.

The research related to the relationship among futures energy prices is conducted by Sadorsky (2000) shows that the futures prices of crude oil, heating oil and unleaded gas are cointegrated with each other. The causality runs from spot energy prices to futures energy prices is in line with Bekiros and Diks (2008). Due to the fact that the futures prices are an indicator that reflects the traders' expectations on future demand conditions (Demirer et al., 2012), this will be reflected from both the spot energy demand as well as futures energy demand.

H3a Spot energy demands are autocorrelated.

H3b Spot energy demands are endogenously time-dependent with future energy prices, spot energy prices, and future energy demand.

The research has illustrated that price would be one of the main factors that will influence the energy demand (Adeyemi & Hunt, 2007; Zhang & Kotani, 2012), therefore, we hypothesize that the spot energy demand will be influenced by both the price level in the spot markets and futures market. As reviewed in the literature sector, there is evidence of price convergence in the spot energy market, and due to the consideration price is one of the main factors influencing the demand level, we hypothesize that the spot energy demand will be correlated with each other. Futures energy prices will significantly influence the level of futures energy demand, while futures energy prices will indirectly exert an impact of spot energy demand through spot energy prices, therefore, we argue that there should be an impact of futures energy demand on spot energy demand. The influence of price on demand in the spot energy market is evidenced by Bonte et al. (2015) who find that the average price elasticity of demand in European Power Exchange spot market is -0.43 between 2010 and 2014.

Table 1 Descriptive statistics

Variables	Minimum	Maximum	Mean	St. Dev	Variation coefficient
Energy (MWh)—spot market	465,578.300	922,465.000	654,659.647	71,056.337	0.109
Energy (MWh)—future market	48.000	1,055,730.000	103,750.166	112,060.387	1.080
Price (Euro/MWh)—spot market	5.779	94.128	49.649	11.455	0.231
Price (Euro/MWh)—future market	11.250	75.148	48.732	6.871	0.141

H4a Future energy demands are autocorrelated.

H4b Future energy demands are endogenously time-dependent with future energy prices, spot energy prices, and spot energy demand.

We hypothesize the autocorrelation among futures energy demand is based on the findings of Lin and Tamvakis (2001) who argue that spillover effects do exist in the futures energy market. As discussed previously, the price level will be an important factor affecting the level of demand in both the spot and futures markets, because there is a level of correlation between spot prices and futures prices, therefore, we argue that there would be a cross-influence of price level on the demand level between spot and futures energy markets. Finally, Silva et al. (2019) find that there is a positive correlation between spot and futures market in the Iberian electricity market, because a higher spot energy demand will increase the spot energy prices, the further positive influence on futures energy prices will impact on the futures energy demand.

Readers should recall that, as long as inertia in the time series itself is produced by the autocorrelation, thus smooth variations are obtained while sudden abrupt changes such as peaks and valleys are avoided, on the other hand, endogeneity reflects the fact that either there would be an unclear cause effect or there would be a simultaneous feedback mechanism, both of which have a probability of temporal dependence.

3 Methodology

The HMM approach developed in this paper is built on the premises that energy prices and consumptions are not-only endogenous but also time dependent. A number of approaches such as VAR, ML, and distributional modelling through copulas are employed in conjunction with the HMM approach as described next.

3.1 The data

Data were collected from the OMIP website (<http://www.omip.pt/Downloads/tabid/104/language/pt-PT/Default.aspx>) 2007 to 2017 for both Iberian countries. While Table 1 reports on the descriptive statistics, Table 2 shows the non-parametric Kendall's Tau correlation coefficients, useful for modelling joint distributions using copulas.

Table 1 indicates that both price and energy consumption are asymmetrically distributed and low dispersed (with the exception of future energy demand), Table 2 shows a relevant correlation pattern not only between energy prices in future and spot markets,

Table 2 Non-parametric Kendall's Tau correlation results

	Future energy	Future price	Spot energy	Spot price
Future Energy (MWh)	1			
Future Price (Euro/MWh)	−0.006103973	1		
Spot Energy (MWh)	−0.064405016	0.070074661	1	
Spot Price (Euro/MWh)	−0.038958766	0.510570069	0.14204931	1

but also between prices and energy in spot markets. Figure 1 sheds some light into the endogenous behavior among these variables: one can easily note that both prices and energy consumption reflect the ongoing feedback processes, for a given time t , that occur between future and spot markets. Besides, prices and energy are auto-correlated, at least between t and $t - 1$.

3.2 Vector auto-regressive (VAR) model

The VAR model is formed by a set of K endogenous variables $y_t = (y_{1t}, \dots, y_{kt}, \dots, y_{Kt})$ for $k = 1, \dots, K$. The VAR(p)-process is then defined as (Pfaff, 2008):

$$= A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t, \quad (1)$$

with A_i are $(K \times K)$ coefficient matrices for $i = 1, \dots, p$ and u_t is a K - dimensional process with $E(u_t) = 0$ and time invariant positive definite covariance matrix $E(u_t u_t^T) = \sum_u$ (white noise). Stability is an important characteristic of a VAR(p)-and can be checked by assessing the characteristic polynomial:

$$\det(I_K - A_1 z - \dots - A_p z^p) \neq 0 \quad \text{for } |z| \leq 1. \quad (2)$$

In this research, the VAR(p)-process is reduced to a VAR(1)-process (Pfaff, 2008):

$$\xi_t = A \xi_{t-1} + v_t, \quad (3)$$

with

$$\xi_t = \begin{bmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}, \quad v_t = \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (4)$$

where the dimensions of the stacked vectors ξ_t and v_t is $(Kp \times 1)$ and the dimension of the matrix A is $(K_p \times K_p)$. Time series plots for energy demand and prices in spot and future markets are given in Fig. 2. Smoothed time series, such as in the case of future prices and spot energy may suggest not only the presence of autocorrelation but also endogeneity with other explanatory variables. Conversely, more nervous time series, such as spot price and future energy may indicate negligible autocorrelation and also a limited feedback process with other variables.

Fig. 1 Endogeneity in prices and energy time series (spot and future markets)

Fig. 2 Time series for energy demands and prices in future and spot markets

3.3 Multivariate copulas

Copulas are a useful tool for modelling multivariate dependence (Yan, 2007) with a special focus on the impact of tail dependence on data analytics and forecasting (Wang & Pham,

Fig. 3 Cullen & Frey graphs for the spot and future energy and demand prices

2012). In this research, the first step towards the multivariate copula modelling was the distributional fit for each one of the price and energy variables. The goodness of fit procedure adopted here observed two parts. The first part consisted of an indicative approach to the best fit by the Cullen & Frey graph, showed in Fig. 3 Cullen and Frey (1999). The second part consisted of a Maximum Likelihood Estimation (MLE) to compute the parameters of the chosen distributions in the first step (Delignette-Muller & Dutang, 2014).

A resample size of 100 was adopted here to generate the yellow dots presented in Fig. 3. Table 3 summarizes the best fitting probability distributions and their estimated parameters for each energy price and demand variables. Results suggest that beta distribution exhibited the best fit for energy demand in future and spot markets, while energy prices in future and spot markets were best represented by a logistic distribution.

3.4 Multivariate HMM with trend

Let's consider j time series $\{(X_{t1}, X_{t2}, \dots, X_{tj}): t=1, \dots, T\}$, which are represented compactly as $\{X_t: t=1, \dots, T\}$. Let's also assume that, conditional on $C(T)=\{C_t: t=1, \dots,$

Table 3 Results for the best distributional fit

Variables	Distribution
Spot energy	Beta (17.65, 7.15)
Future energy	Beta (0.77, 6.49)
Spot price	Logistic (49.46, 6.40)
Future price	Logistic (48.49, 3.69)

$T\}$, these random vectors are mutually independent (Zucchini et al., 2016). To specify an HMM for such time series it is necessary to postulate a model for the distribution of the random vector X_t in each of the m states of the parameter process. That is, one requires the following probabilities to be specified for $t = 1, 2, \dots, T$, $i = 1, 2, \dots, m$, and all relevant x : $\text{tpi}(x) = P(X_t = x | C_t = i)$. Let's keep the time index t , thus allowing the state dependent probabilities to change over time. It should be noted that it is not required that each of the time series have the same distribution.

In this research, we use multivariate copulas, as discussed in the previous section, to model stochastic variations in HMM. Once the required joint distributions have been selected, i.e. state-dependent probabilities $\text{tpi}(x_t)$ have been specified, the likelihood of a general multivariate HMM has the same form as that of basic HMM models:

$$L_T = \delta_1 P(X_1) \Gamma_2 P(X_2) \dots \Gamma_T P(X_T) 1', \quad (5)$$

where $X_1, \dots, X_T$ are the observations and the likelihood function is given as (Zucchini et al., 2016):

$${}_t P(X_t) = \text{diag}({}_t p_1(X_t), \dots, {}_t p_m(X_t)) \quad (6)$$

Equations (5) and (6) hold for continuous-valued series, provided that probabilities are interpreted as densities where necessary. Therefore, in the stochastic multivariate HMM model adopted in this research, VAR and ML models are used in a complementary fashion. In fact, the model $\text{tpij} = a_{ij} + b_t$ is taken as the starting point for the pseudo-code further discussed, where tpij is the probability that time series j belongs to state m on day t , observing the following set definition given in the first paragraph of this section:

$T = \{1 \dots \text{days}\}$, where days is the number of days of the sample;

$J = \{\text{spot energy prices (sp), future energy prices (fp), spot energy demand (se), future energy demand (fe)}\}$;

$M = \{m1 \dots m256\}$, where each m is given observing the quantile combinations for each individual time series of energy prices and demands, as given in Table 4.

4 Presentation and discussion of results

Figures 4 and 5 present the results of transition matrix, which includes fifteen more portable states and their respective subsets. we can see that there are 235 individual states out of 256 have more than zero probabilities, however, there are only 1580 possible combinations out of 55225 have more than zero probabilities that the results indicate that, when using Markov chains and their transition probabilities to analyze energy demands and prices, there is only a small proportion of possible movements. This can be possibly attributed to the strong feedback mechanisms between energy prices and demands in future and spot

Table 4 Multivariate HMM states

State (n)	Sp	fp	se	fe
m1	abQ3	abQ3	abQ3	abQ3
m2	btMQ3	abQ3	abQ3	abQ3
...	...	...	...	...
m255	btQ1M	beQ1	beQ1	beQ1
m256	beQ1	beQ1	beQ1	beQ1

Legend: abQ3—above quantile 3; beQ1—below quantile 1; btQ1M—between quantile 1 and median; btMQ3—between median and quantile 3

Fig. 4 The most probable states' Pareto plot

markets in parallel with autocorrelation that reduces the level of abrupt fluctuations. With regard to future and spot energy prices, prices are found to be presented in the same quantile for the 15 more probable individual states (cf. Table 4, footnote), while future and spot energy demands present a more detached behavior. This maybe the result of prospective and speculative movements with respect to future energy demand behavior. The detached behaviour between spot energy demand and futures energy demand can be partly attributed to the fact that investment in the spot market has a lower level of risk, the information can be more easily obtained compared to the investment in the futures market. In comparison, investment in the futures markets require a larger volume of funds and also it suffers from a higher level of risk, it requires the investors to have a certain level of skill and experience in the investment activities, because these two investments have different characteristics and feature, this determines the differentiated behaviour of demands in these two markets. In reality, Fig. 5 shows the probable movement between different states, a subtle movement in the quantile interval of future energy triggers the transition, all other three variables

Fig. 5 Most probable state transitions' heat map (zero probability is represented by blank cells)

remaining at their respective quantiles. It is also interesting to note that the two more probable movements are related to transitions to the same state, thus suggesting inertia.

The convergence of spot prices in the current study is in accordance with the finding of King and Cuc (1996). This is in line with our hypothesis H1a. While the detached behaviour between spot energy demand and futures energy demand seems not be explicitly addressed by the previous literature yet. This investigation of this issue is very important due to the fact that understanding the demand behaviour in the spot and futures markets will provide a basis to further understand and formulate a reasonable price mechanism for the market. We further find that the in general spot prices and futures price are related with each other, this is also in line with the previous literature as reviewed in Sect. 2.

The results of the bootstrapped VAR model as well as the feedback mechanisms among energy price and demand in the spot and futures markets are reported in Figs. 6 and 7, respectively. In terms of Fig. 6, if the horizontal solid line in zero does not cross the dispersion diagram, that indicates that the variable is significant. We further find that the significant relationships between energy price and energy demand are positive. As reflected from both these two figures, there is a complex feedback mechanism between energy price and demand in both the spot and futures markets. In particular, Fig. 7 is the response to the hypotheses raised in the end of Sect. 2.

It is interesting to note that, as suspected in Fig. 2, spot prices and future energy are not autocorrelated, thus implying smaller inertia in their fluctuations. As a matter of fact, spot prices are only time dependent on future prices and spot energy, this result partly supports our hypothesis H1b, however, our results are not exactly in line with the argument of Pindyck (2001), we find that the spot prices are not time dependent on future energy demand. While future energy is solely time dependent on spot energy behavior. These results suggest that the behaviour of futures energy demand is affected by the spot energy demand not the important factor futures energy prices as discussed previously. This indicates that in the Iberian energy markets, the investors normally focus more on the spot price movement in their investment decisions in the futures market.

On the other hand, future prices are not only autocorrelated but also time-dependent with spot energy and future energy demands level, thus suggesting that the price formation mechanisms differ between spot and future prices. The autocorrelation of future prices is in

Fig. 6 The bootstrapped VAR with 2 lagged periods' coefficients

line with Lin and Tamvakis (2001) and in accordance with our hypothesis H2a. The findings also partly support our hypothesis H2b, however, the findings are different from Bekiros and Diks (2008) who report that there is a causality running from spot prices to future prices, this difference is mainly attributed to the fact that we adopt a different methodology and different data period was used. This has a number of policy implications for the Iberian energy industry, for instance, capacity expansion should be one of the focuses in the Iberian energy sector through introducing additional investment. This additional supply of energy in the market will not only increase the competitive conditions but also stabilize the price level on the futures market and reduce the risk for the companies purchasing the futures contract.

A similar behavior occurs with spot energy, which is autocorrelated and time-dependent with future prices and future energy. Results suggest that part of the behavior of spot energy time series is not only driven by current demand levels, but also by recent expectations on future behavior of energy and future prices. One of the scenarios is a higher level of futures price will induce an increase in the spot energy demand. We can also see from our results that the spot energy and futures energy demands are correlated with each other. This result is partly in line with our hypotheses H3b and H4b. this finding also accord with Silva et al. (2019).

Finally, Figs. 8 and 9 presents the results from the bootstrapped multinomial logit regression, which provides more information about the underlying more probable states between energy price and energy demand. We notice that there is a level of stability of the probabilities of belonging to a given state although there are few exceptions (cf. Figure 8), in addition, the figure shows that the significant trend components do not influence these states (cf. Figure 9), this finding can also be supported by Fig. 2. These results suggest that there is no violation on the main assumptions of HMM with regard to its application to the Iberian energy industry.

Fig. 7 Feedback mechanisms between energy prices and demands in future and spot markets

Different from the previous literature, which finds that the price elasticity of demand is negative, indicating that there would be a negative relationship between price and demand level, the current study finds that the relationship is significant and positive, suggesting that in the Iberian energy industry, there is a large volume of demand without any influence by the price level, therefore, one of the recommendations to the Iberian energy market is to encourage investment in this energy sector, expand the energy producing capacity, which would be helpful to form a reasonable market trading mechanism.

Fig. 8 Results on bootstrapped multinomial logit model for the probability of 15 more probable individual states

5 Conclusions

The energy sector is very important for the economic and societal development for all countries in the world, and for the Iberian energy industry specifically, it suffers from a number of issues. For the Spanish energy sector, numerous sources of uncertainties arise because of the historically highly unpredictability of the sector itself. For the Portuguese case, a balancing mechanism is urgently needed to implement renewable resources and encourage new investment through providing more capacity incentives (Amorim et al., 2013). In terms of two specific trading patterns in the energy market, the spot market and future market, Ciarreta and Zarraga (2016) argue that The Iberian energy sector has the characteristics of price and transmission volatility, while Capitán-Herráiz and Rodríguez-Monroy (2013) find that the Iberian energy sector lacks of liquidity in the futures energy market. These issues can be solved thorough different ways from various perspectives, while one possible route is to provide more detailed and depth analysis about the price mechanism in the spot and futures energy markets by investigating the inter-temporary interactions between the price level and the demand conditions. In addition, for both of these two countries, sustainable economic development has been an issue concerned by the governments following the hard hit of the financial crisis, investment in the renewable energy will not only promote sustainable development, increase the employment and improve the economic growth, but also will reduce the energy external dependency of these two countries. Therefore, the process of decarbonization is the focus of these two countries in their process of moving toward a greener economy. Investigating the inter-temporal relationship between demand and price level in the spot and futures energy market will be helpful not only to establish a marketable price mechanism, but will also provide more guidance, information and motivation for these two countries to further increase their effort in reducing the level of pollution for energy generation.

The current study significantly contributes to the literature in terms of spot and futures energy markets by empirically investigating the inter-connections among spot energy prices, spot energy demand, futures energy prices as well as futures energy demand. Methodologically, the current study proposed a stochastic Hidden Markov Model, which benefits from the advantage of being able to examine the endogeneity and temporal dependence under the facilitation of bootstrapped vector auto-regressive model and multinomial logit model. Our findings suggest that when using Markov chains and their transition probabilities to analyze energy demand and prices, there is only a small proportion of possible

Fig. 9 results of bootstrapped multinomial logit model on the trend coefficients of 15 most probable states

movements; in addition, we find that all significant relationships between energy prices and demands are positive; it is shown that spot prices are only time dependent on future prices and spot energy, while future energy is solely time dependent on spot energy behavior; the results further report that future prices are not only autocorrelated but also time-dependent with spot energy and future energy demands level; finally, we find that spot energy is auto-correlated and time-dependent with future prices and future energy.

Our study generates policy implications to solve the issues faced by the Iberian energy market. In order to improve the liquidity in the futures market, the price setting up mechanism can be based on the spot and futures energy demand. This is supposed to set up a price level that will more accurately reflect the demand conditions and further improve the level of liquidity (the seller would be able to sell the energy in the futures markets more quickly). On the other hand, the positive relationship between energy price and energy demand indicates that higher energy price leads to an increase in the energy demand, this can be understood from the perspective that energy is a necessity product, the price mechanism can not only reflect the demand condition, but it can be used to adjust the level of energy demand. In other words, different types of policies can be used during different time (i.e. during the period of energy supply shortage, the price level can be reduced to decrease energy demand, in comparison, during the time of energy supply surplus, the energy price can be lifted up). Obviously, the volatility in the price level of the spot market can be alleviated by considering the level in the spot and futures energy demand, while probably increasing the incentive for the incumbent firms to invest in the energy sector will not only increase the competitive conditions in the Iberian energy sector considering that there are only few suppliers in the market, but also plays an important role in stabilize the price level on the spot market. The main purpose of the futures energy market is to reduce the price volatility and future risk, from our results, it shows that spot energy demand and future energy prices and futures energy demand are time dependent with each other, therefore, in order to reduce the price volatility in the energy futures market, the priority is to increase the stability in the spot and futures energy demand.

As discussed in the introduction, for the future research, we would make use of our proposed HMM model to predict and assess the decarbonized alternative sources of energy with the facilitation of financial equations.¹

¹ The formulas for the calculation of LACE and LCOE are not reported in this paper because of space and word limit, but it would be available upon request from the corresponding author.

Appendix: Full R Code

```

#Clean memory
rm(list = ls())
gc()

#Librarys

library(copula)
library(ggplot2)
library(grid)
library(igraph)
library(gtools)
library(fitdistrplus)
library(hmm.discnp)
library(vars)
library(nnet)
library(corrgram)

#Functions

#A modification of function descdist to remove the graph's main
descdist.2 <- function(data, discrete = FALSE, boot = NULL, method = "unbiased", graph =
TRUE,
                        obs.col = "darkblue", obs.pch = 16, boot.col = "orange")
{
  #if(is.mcnode(data)) data <- as.vector(data)
  if (missing(data) || !is.vector(data,mode="numeric"))
    stop("data must be a numeric vector")
  if (length(data) < 4)
    stop("data must be a numeric vector containing at least four values")
  moment <- function(data,k){
    m1 <- mean(data)
    return(sum((data-m1)^k)/length(data))
  }
  if (method=="unbiased")
  {
    skewness <- function(data){
      # unbiased estimation (Fisher 1930)
      sd <- sqrt(moment(data,2))
      n <- length(data)
      gamma1 <- moment(data,3)/sd^3
      unbiased.skewness <- sqrt(n*(n-1)) * gamma1 / (n-2)
      return(unbiased.skewness)
    }
    kurtosis <- function(data){
      # unbiased estimation (Fisher 1930)
      n <- length(data)
      var<-moment(data,2)
      gamma2 <- moment(data,4)/var^2
      unbiased.kurtosis <- (n-1)/((n-2)*(n-3)) * ((n+1)*gamma2 -3*(n-1)) + 3
      return(unbiased.kurtosis)
    }
    standdev <- function(data){
      sd(data)

```

```

    }
  }
else
  if (method=="sample")
  {
    skewness <- function(data)
    {
      sd <- sqrt(moment(data, 2))
      return(moment(data, 3)/sd^3)
    }
    kurtosis <- function(data)
    {
      var <- moment(data, 2)
      return(moment(data, 4)/var^2)
    }
    standdev <- function(data){
      sqrt(moment(data,2))
    }
  }
else
  stop("The only possible value for the argument method are 'unbiased' or 'sample'")

res <- list(min=min(data),max=max(data),median=median(data),
           mean=mean(data),sd=standdev(data),
           skewness=skewness(data),kurtosis=kurtosis(data), method = method)

skewdata<-res$skewness
kurtdata<-res$kurtosis

# Cullen and Frey graph
if (graph) {
  # bootstrap sample for observed distribution
  # and computation of kurtmax from this sample
  if (!is.null(boot)) {
    if (!is.numeric(boot) || boot<10) {
      stop("boot must be NULL or a integer above 10")
    }
    n<-length(data)

    databoot<-matrix(sample(data,size=n*boot,replace=TRUE),nrow=n,ncol=boot)
    s2boot<-sapply(1:boot,function(iter) skewness(databoot[,iter])^2)
    kurtboot<-sapply(1:boot,function(iter) kurtosis(databoot[,iter]))

    kurtmax<-max(10,ceiling(max(kurtboot)))
    xmax<-max(4,ceiling(max(s2boot)))
  }
  else{
    kurtmax<-max(10,ceiling(kurtdata))
    xmax<-max(4,ceiling(skewdata^2))
  }

  ymax<-kurtmax-1
  plot(skewdata^2,kurtmax-kurtdata,pch=obs.pch,xlim=c(0,xmax),ylim=c(0,ymax),
       yaxt="n",xlab="square of skewness",ylab="kurtosis",main="")

```

```

yax<-as.character(kurtmax-0:ymax)
axis(side=2,at=0:ymax,labels=yax)
if (!discrete) {
  # beta dist
  p<-exp(-100)
  lq<-seq(-100,100,0.1)
  q<-exp(lq)
  s2a<-(4*(q-p)^2*(p+q+1))/((p+q+2)^2*p*q)
  ya<-kurtmax-(3*(p+q+1)*(p*q*(p+q-6)+2*(p+q)^2)/(p*q*(p+q+2)*(p+q+3)))
  p<-exp(100)
  lq<-seq(-100,100,0.1)
  q<-exp(lq)
  s2b<-(4*(q-p)^2*(p+q+1))/((p+q+2)^2*p*q)
  yb<-kurtmax-(3*(p+q+1)*(p*q*(p+q-6)+2*(p+q)^2)/(p*q*(p+q+2)*(p+q+3)))
  s2<-c(s2a,s2b)
  y<-c(ya,yb)
  polygon(s2,y,col="lightgrey",border="lightgrey")
  # gamma dist
  lshape<-seq(-100,100,0.1)
  shape<-exp(lshape)
  s2<-4/shape
  y<-kurtmax-(3+6/shape)
  lines(s2,y,lty=2)
  # lnorm dist
  lshape<-seq(-100,100,0.1)
  shape<-exp(lshape)
  es2<-exp(shape^2)
  s2<-(es2+2)^2*(es2-1)
  y<-kurtmax-(es2^4+2*es2^3+3*es2^2-3)
  lines(s2,y,lty=3)

  legend(xmax*0.2,ymax*1.03,pch=obs.pch,legend="Observation",bty="n",cex=0.8,pt.cex=1.2,col=obs.col)
  if (!is.null(boot)) {
    legend(xmax*0.2,ymax*0.98,pch=1,legend="bootstrapped values",
           bty="n",cex=0.8,col=boot.col)
  }
  legend(xmax*0.55,ymax*1.03,legend="Theoretical distributions",bty="n",cex=0.8)
  legend(xmax*0.6,0.98*ymax,pch=8,legend="normal",bty="n",cex=0.8)
  legend(xmax*0.6,0.94*ymax,pch=2,legend="uniform",bty="n",cex=0.8)
  legend(xmax*0.6,0.90*ymax,pch=7,legend="exponential",bty="n",cex=0.8)
  legend(xmax*0.6,0.86*ymax,pch=3,legend="logistic",bty="n",cex=0.8)
  legend(xmax*0.6,0.82*ymax,fill="grey80",legend="beta",bty="n",cex=0.8)
  legend(xmax*0.6,0.78*ymax,lty=3,legend="lognormal",bty="n",cex=0.8)
  legend(xmax*0.6,0.74*ymax,lty=2,legend="gamma",bty="n",cex=0.8)
  legend(xmax*0.58,0.69*ymax,legend=c("(Weibull is close to gamma and lognormal)"),
        bty="n",cex=0.6)
}
else {
  # negbin dist
  p<-exp(-10)
  lr<-seq(-100,100,0.1)
  r<-exp(lr)
  s2a<-(2-p)^2/(r*(1-p))

```

```

ya<-kurtmax-(3+6/r+p^2/(r*(1-p)))
p<-1-exp(-10)
lr<-seq(100,-100,-0.1)
r<-exp(lr)
s2b<-(2-p)^2/(r*(1-p))
yb<-kurtmax-(3+6/r+p^2/(r*(1-p)))
s2<-c(s2a,s2b)
y<-c(ya,yb)
polygon(s2,y,col="grey80",border="grey80")

legend(xmax*0.2,ymax*1.03,pch=obs.pch,legend="Observation",bty="n",cex=0.8,pt.cex=1.2,
col = obs.col)
if (!is.null(boot)) {
  legend(xmax*0.2,ymax*0.98,pch=1,legend="bootstrapped values",
    bty="n",cex=0.8,col=boot.col)
}
legend(xmax*0.55,ymax*1.03,legend="Theoretical distributions",bty="n",cex=0.8)
legend(xmax*0.6,0.98*ymax,pch=8,legend="normal",bty="n",cex=0.8)
legend(xmax*0.6,0.94*ymax,fill="grey80",legend="negative binomial",
  bty="n",cex=0.8)
legend(xmax*0.6,0.90*ymax,lty=2,legend="Poisson",bty="n",cex=0.8)
# poisson dist
llambda<-seq(-100,100,0.1)
lambda<-exp(llambda)
s2<-1/lambda
y<-kurtmax-(3+1/lambda)
lines(s2,y,lty=2)
}
# bootstrap sample for observed distribution
if (!is.null(boot)) {
  points(s2boot,kurtmax-kurtboot,pch=1,col=boot.col,cex=0.5)
}
# observed distribution
points(skewness(data)^2,kurtmax-kurtosis(data),pch=obs.pch,cex=2,col=obs.col)
# norm dist
points(0,kurtmax-3,pch=8,cex=1.5,lwd=2)
if (!discrete) {
  # unif dist
  points(0,kurtmax-9/5,pch=2,cex=1.5,lwd=2)
  # exp dist
  points(2^2,kurtmax-9,pch=7,cex=1.5,lwd=2)
  # logistic dist
  points(0,kurtmax-4.2,pch=3,cex=1.5,lwd=2)
}
} # end of is (graph)

return(structure(res, class = "descdist"))

}

#Function with states names and specs
make.specs <- function(){
  #state vector
  states <- c("abQ3", "btMQ3", "btQ1M", "beQ1")

```

```

vet.state <- data.frame(matrix(NA, ncol = 5, nrow = 256))
pos <- 1
for(i in 1:4){
  for(j in 1:4){
    for(k in 1:4){
      for(l in 1:4){
        vet.state[pos, 2] <- states[l]
        vet.state[pos, 3] <- states[k]
        vet.state[pos, 4] <- states[j]
        vet.state[pos, 5] <- states[i]
        pos <- pos + 1
      }
    }
  }
}
vet.state[,1] <- paste0("m", as.character(1:256))
names(vet.state) <- c("State", "Spot Price", "Future Price", "Spot Energy", "Future Energy")
return(vet.state)
}

#Function to indentify the state M from a matrix J
make.state <- function(x){
  #beQ1 – below quantile 1;
  #btQ1M – between quantile 1 and median;
  #btMQ3 – between median and quantile 3;
  #abQ3 – above quantile 3;

  #Backup x matrix
  x.back <- x
  #matrix J
  base.state <- x
  for(i in 1:dim(base.state)[2]){
    val.qnt <- quantile(base.state[,i])
    aux <- data.frame(ifelse(base.state[,i] < val.qnt[2], "beQ1",
                           ifelse(base.state[,i] < val.qnt[3], "btQ1M",
                           ifelse(base.state[,i] < val.qnt[4], "btMQ3", "abQ3"))))
    base.state[,i] <- aux
  }
  #converte into a string
  for(i in 1:dim(base.state)[2]){
    base.state[,i] <- as.character(base.state[,i])
  }

  #matrix with states names and specs
  vet.state <- make.specs()

  vet.state.aux <- NULL
  for(i in 1:dim(vet.state)[1]){
    str.aux <- vet.state[i,2]
    for(j in 3:5){
      str.aux <- paste(str.aux, vet.state[i,j], sep = "-")
    }
    vet.state.aux <- rbind(vet.state.aux, c(vet.state[i,1], str.aux))
  }
}

```



```

base.state.aux <- NULL
for(i in 1:dim(base.state)[1]){
  str.aux <- base.state[i,1]
  for(j in 2:4){
    str.aux <- paste(str.aux, base.state[i,j], sep = "-")
  }
  base.state.aux <- rbind(base.state.aux, str.aux)
}

states.aux <- NULL
for(i in 1:dim(base.state)[1]){
  states.aux <- c(states.aux, which(vet.state.aux[,2] %in% as.character(base.state.aux[i,])))
}

final.states <- as.data.frame(vet.state.aux[states.aux,1])
names(final.states) <- "States"

return(final.states)
}

#function to plot a heatmap of most prob. changes
plot.heat <- function(x, val, val.most){
  Mt.graph <- data.frame(matrix(NA, ncol = 3, nrow = dim(Mt)[1] ** 2))
  for(i in 1:dim(x)[1]){
    ini <- (i-1)*dim(x)[1] + 1
    fim <- i*dim(Mt)[1]

    Mt.graph[ini:fim, 1] <- names(x)[i]
    Mt.graph[ini:fim, 2] <- names(x)
    Mt.graph[ini:fim, 3] <- as.numeric(x[,i])
  }
  Mt.graph <- Mt.graph[Mt.graph[,3] != 0,]
  names(Mt.graph) <- c("Origin", "Destiny", "Prob")

  Mt.graph <- Mt.graph[order(Mt.graph[,3], decreasing = TRUE)[1:val],]

  #most prob states to plot
  Most <- Mt.graph[1:val.most,1:2]
  most.states <- c(Most[,1],Most[,2])
  most.states <- mixedsort(levels(factor(most.states)), decreasing = TRUE)

  #Remake the matrix for plot
  x.dim <- length(levels(factor(Mt.graph[,1])))
  y.dim <- length(levels(factor(Mt.graph[,2])))
  Mt.graph2 <- data.frame(matrix(0, ncol = 3, nrow = x.dim*y.dim))

  for(i in 1:length(mixedsort(levels(factor(Mt.graph[,1]))))){
    ini <- (i-1)*length(mixedsort(levels(factor(Mt.graph[,2])))) + 1
    fim <- i*length(mixedsort(levels(factor(Mt.graph[,2]))))

    Mt.graph2[ini:fim, 1] <- mixedsort(levels(factor(Mt.graph[,1])))[i]
    Mt.graph2[ini:fim, 2] <- mixedsort(levels(factor(Mt.graph[,2])))[i]
  }
  names(Mt.graph2) <- names(Mt.graph)
  for(i in 1:dim(Mt.graph2)[1]){

```

```

Mt.graph.aux <- Mt.graph[Mt.graph[,1] == Mt.graph2[i,1],]
Mt.graph.aux <- Mt.graph.aux[Mt.graph.aux[,2] == Mt.graph2[i,2],]
Mt.graph2[i,3] <- ifelse(dim(Mt.graph.aux)[1] == 0, 0, as.numeric(Mt.graph.aux[,3]))
}
Mt.graph2[,1] <- factor(Mt.graph2[,1], levels = mixedsort(levels(factor(Mt.graph2[,1]))))
Mt.graph2[,2] <- factor(Mt.graph2[,2], levels = mixedsort(levels(factor(Mt.graph2[,2]))))
Mt.graph2[,3] <- Mt.graph2[,3]*100 #convert to %

specs.states <- make.specs()
specs.states <- specs.states[sort(which(specs.states[,1] %in% most.states), decreasing =
TRUE),]

ggplot(data = Mt.graph2) +
  geom_tile(mapping = aes(x = Origin, y = Destiny, fill = Prob)) +
  scale_fill_gradient(low = "white", high = "darkblue") +
  ylab("Destiny State") +
  xlab("Origin State") +

  annotate(geom = "text", x = rep((1+x.dim), length(most.states)), y = (y.dim-
(1:length(most.states))),
    label = rep("", length(most.states))) +
  annotate(geom = "text", x = rep((2+x.dim), 1+length(most.states)), y = (y.dim-
(1:(1+length(most.states)))),
    label = c(names(specs.states)[1], as.character(specs.states[,1]))) +
  annotate(geom = "text", x = rep((3+x.dim), 1+length(most.states)), y = (y.dim-
(1:(1+length(most.states)))),
    label = rep("-", 1+length(most.states))) +
  annotate(geom = "text", x = rep((4+x.dim), 1+length(most.states)), y = (y.dim-
(1:(1+length(most.states)))),
    label = c(names(specs.states)[2], as.character(specs.states[,2]))) +
  annotate(geom = "text", x = rep((6+x.dim), 1+length(most.states)), y = (y.dim-
(1:(1+length(most.states)))),
    label = c(names(specs.states)[3], as.character(specs.states[,3]))) +
  annotate(geom = "text", x = rep((8+x.dim), 1+length(most.states)), y = (y.dim-
(1:(1+length(most.states)))),
    label = c(names(specs.states)[4], as.character(specs.states[,4]))) +
  annotate(geom = "text", x = rep((10+x.dim), 1+length(most.states)), y = (y.dim-
(1:(1+length(most.states)))),
    label = c(names(specs.states)[5], as.character(specs.states[,5]))) +
  annotate(geom = "text", x = rep((11+x.dim), length(most.states)), y = (y.dim-
(1:length(most.states))),
    label = rep("", length(most.states)))
}

#function to plot coef. from var
plot.coef <- function(x, ylb){
  mat.coef <- x
  names.var <- names(mat.coef)

  Gr <- as.data.frame(matrix(NA, nrow=(length(names.var)*2), ncol = 4))
  names(Gr) <- c("Context", "Var", "Boot_95_5", "Est")
  for(ctx in 1:length(names.var)){#Por contextual

```

```

pos <- (ctx-1)*2 + 1

#Contextual
Gr[pos:(pos+1),1] <- names.var[ctx]

#Var
Gr[pos:(pos+1),2] <- "Variable"

#Combined Forecast
#Estimate
Gr[pos+1,3]<-quantile(mat.coef[,ctx], 0.05)
Gr[pos,3]<-quantile(mat.coef[,ctx], 0.95)

Gr[pos,4]<-median(mat.coef[,ctx])
}

Gr <- within(Gr, Context <- factor(Context,levels=names.var))

ggplot() +
  facet_wrap(~Context, nrow = 1, scales = "free", strip.position = "bottom") +
  geom_hline(yintercept = 0, lwd=1.0) + #Linha horizontal que cruza o y=0
  geom_point(data = Gr, aes(x=Context, y=Est), shape = 1, size = 6.5) +
  geom_line(data = Gr, aes(x=Context, y=Boot_95_5),lwd = 0.75, lty="dashed") +
  ylab(ylb) +
  theme(panel.background = element_rect(fill='white', colour='gray'),
        axis.title.x=element_blank(), axis.title.y=element_text(size = 12,
margin=margin(0,15,0,0)), text = element_text(size = 12),
        axis.text.x = element_blank(),
        strip.background = element_blank())
}

#plot coefficients for multinomial logistic regression
plot.coef.mlr <- function(x, context, name.title = ""){
  aux <- x
  names.context <- context

  GrContext01 <- as.data.frame(matrix(NA, nrow =(length(names.context)*2) ,ncol = 4))
  names(GrContext01) <- c("Context", "State", "Boot_95_5", "Est")
  for(ctx in 1:length(names.context)){#Por contextual
    pos <- (ctx-1)*2 + 1

    #Contextual
    GrContext01[pos:(pos+1),1] <- names.context[ctx]

    #State
    GrContext01[pos:(pos+1),2] <- as.character(aux[1,1])

    #Combined Forecast
    #Estimate
    GrContext01[pos+1,3]<-quantile(as.numeric(as.character(aux[,ctx+2])), 0.05)
    GrContext01[pos,3]<-quantile(as.numeric(as.character(aux[,ctx+2])), 0.95)

    GrContext01[pos,4]<-median(as.numeric(as.character(aux[,ctx+2])))
  }
}

```

```

GrContext01 <- within(GrContext01, Context <- factor(Context,levels=names.context))

GrC <- GrContext01

setwd(path_results)
ggplot() +
  facet_wrap(~Context, nrow = 1, scales = "free") +
  geom_hline(yintercept = 0, lwd=1.0) + #Linha horizontal que cruza o y=0
  geom_point(data = GrC, aes(x=Context, y=Est), shape = 1, size = 6.5) +
  geom_line(data = GrC, aes(x=Context, y=Boot_95_5),lwd = 0.75, lty="dashed") +
  ylab(GrC[1,2]) +
  labs(title = name.title) +
  theme(axis.title.x=element_blank(),
        strip.background = element_blank(),
        strip.text.x = element_blank())
}

#Paths
#set the paths
path_base<-"/home/jorge/Dropbox/A CODIGOS DEA MURILO/Backup Computador
Jorge/Iberian Energy HMM/Bases/"
path_results<-"/home/jorge/Dropbox/A CODIGOS DEA MURILO/Backup Computador
Jorge/Iberian Energy HMM/Results/"

#Read Data
setwd(path_base)

base <- read.csv(file = "Base.csv", header = TRUE, stringsAsFactors = FALSE, fileEncoding =
"UTF-8")
names.base <- as.character(read.csv(file = "Base.csv", header = FALSE, stringsAsFactors =
FALSE, fileEncoding = "UTF-8")[1,])
names(base) <- names.base

#Base
base <- cbind(base, 1:dim(base)[1], (1:dim(base)[1])*(1:dim(base)[1]))
base <- base[,c(1,5,3,4,2,6,7)]
base[,1] <- as.Date(base[,1])
names(base)[6:7] <- c("Trend", "Trend2")

#Kendall's correlation matrix
setwd(path_results)
capture.output(cor(base[,2:5], method = "kendall"), file = "Kendall cor.txt", append = FALSE)

png(filename = "Correlogram.png", width = 6, height = 6, units = "in", res = 72)
corrgram(x = base[,2:5], cor.method = "kendall", lower.panel = panel.pie, main = "Kendall")
dev.off()

#Scatterplots and temporal series plots
setwd(path_results)

#1 - spot price [t] x spot energy [t]
#2 - future price [t] x future energy [t]
#3 - spot price [t] x future price [t]
#4 - spot energy [t] x future energy [t]
#5 - spot price [t] x spot price [t-1]

```

```

#6 - future price [t] x future price [t-1]
#7 - spot energy [t] x spot energy [t-1]
#8 - future energy [t] x future energy [t-1]

png(filename = "Scatter.png", width = 9, height = 15, units = "in", res = 144)
par(mfrow = c(4,2))
#1 - spot price [t] x spot energy [t]
plot(x = base[,4], y = base[,5],
     main = "Spot Price [t] x Spot Energy [t]",
     xlab = "Spot Energy [MWh]", ylab = "Spot Price [€/MWh]")
#2 - future price [t] x future energy [t]
plot(x = base[,2], y = base[,3],
     main = "Future Price [t] x Future Energy [t]",
     xlab = "Future Energy [MWh]", ylab = "Future Price [€/MWh]")
#3 - spot price [t] x future price [t]
plot(x = base[,5], y = base[,3],
     main = "Spot Price [t] x Future Price [t]",
     xlab = "Spot Price [€/MWh]", ylab = "Future Price [€/MWh]")
#4 - spot energy [t] x future energy [t]
plot(x = base[,4], y = base[,2],
     main = "Spot Energy [t] x Future Energy [t]",
     xlab = "Spot Energy [MWh]", ylab = "Future Energy [MWh]")
#5 - spot price [t] x spot price [t-1]
plot(y = base[(2:dim(base)[1]),5], x = base[(1:(dim(base)[1]-1)),5],
     main = "Spot Price [t] x Spot Price [t-1]",
     ylab = "Spot Price (t) [€/MWh]", xlab = "Spot Price (t-1) [€/MWh]")
#6 - future price [t] x future price [t-1]
plot(y = base[(2:dim(base)[1]),3], x = base[(1:(dim(base)[1]-1)),3],
     main = "Future Price [t] x Future Price [t-1]",
     ylab = "Future Price (t) [€/MWh]", xlab = "Future Price (t-1) [€/MWh]")
#7 - spot energy [t] x spot energy [t-1]
plot(y = base[(2:dim(base)[1]),4], x = base[(1:(dim(base)[1]-1)),4],
     main = "Spot Energy [t] x Spot Energy [t-1]",
     ylab = "Spot Energy (t) [MWh]", xlab = "Spot Energy (t-1) [MWh]")
#8 - future energy [t] x future energy [t-1]
plot(y = base[(2:dim(base)[1]),2], x = base[(1:(dim(base)[1]-1)),2],
     main = "Future Energy [t] x Future Energy [t-1]",
     ylab = "Future Energy (t) [MWh]", xlab = "Future Energy (t-1) [MWh]")

dev.off()

#Temporal Series
png(filename = "Temporal Series.png", width = 8, height = 14, units = "in", res = 144)
par(mfrow = c(2,1))
plot(x = base[,1], y = base[,5], type = "l", col = 2,
     xlab = "Year", ylab = "Price [€/MWh]", main = "Prices Series",
     ylim = c(min(base[,5],base[,3]),max(base[,5],base[,3])))
lines(x = base[,1], y = base[,3], col = 3)
legend(x = "topleft", legend = c("Spot Price", "Future Price"), col = 2:3, lwd = 1)
plot(x = base[,1], y = base[,4], type = "l", col = 2,
     xlab = "Year", ylab = "Energy [MWh]", main = "Energy Series",
     ylim = c(min(base[,4],base[,2]),max(base[,4],base[,2])))
lines(x = base[,1], y = base[,2], col = 3)
legend(x = "topleft", legend = c("Spot Energy", "Future Energy"), col = 2:3, lwd = 1)

```

```

dev.off()

#States
#beQ1 – below quantile 1;
#btQ1M – between quantile 1 and median;
#btMQ3 – between median and quantile 3;
#abQ3 – above quantile 3;

#matrix with states
vet.state <- make.specs()

#identify the state from J matrix
base <- cbind(base[,1], make.state(base[,2:5]), base[,2:7])
names(base)[1:2] <- c("Day", "State")

#Prob. Matrix
#prob. matrix for state m
M <- as.data.frame(vet.state[,1])
prob.aux <- NULL
for(i in 1:dim(M)[1]){
  prob.aux <- c(prob.aux, length(which(base[,2] %in% M[i,1])))
}
prob.aux <- prob.aux/sum(prob.aux)
M <- as.data.frame(cbind(M, as.numeric(prob.aux)))

#remove prob = 0
M <- M[(M[,2] != 0), ]

names(M) <- c("States", "Prob")

#Transition Matrix
Mt <- data.frame(matrix(0, ncol = dim(M)[1], nrow = dim(M)[1]))
names(Mt) <- M[,1]
rownames(Mt) <- M[,1]

for(i in 1:dim(Mt)[1]){
  #index for state i
  idx <- which(base[,2] %in% as.character(M[i,1]))
  #index for state i+1
  idx.tr <- idx + 1
  idx.tr <- as.numeric(na.omit(ifelse(idx.tr == (dim(base)[1]+1), NA, idx.tr)))

  M.aux <- data.frame(as.character(base[idx.tr,2]))
  M.aux.split <- split(M.aux, M.aux[,1])
  for(j in 1:length(M.aux.split)){
    #index in Mt matrix
    idx.mt <- which(M[,1] %in% as.character(M.aux.split[[j]][1,1]))
    Mt[i,idx.mt] <- length(M.aux.split[[j]][,1])
  }
}

#Normalize Mt
Mt <- Mt/sum(Mt)

#pareto plot for M

```

```

M.ordered <- M[order(M[,2], decreasing = TRUE), ]
M.ordered[,2] <- M.ordered[,2]*100
M.ordered <- M.ordered[1:15,]
M.ordered <- as.data.frame(cbind(as.character(M.ordered[,1]), 1:15, M.ordered[,2],
cumsum(M.ordered[,2])))
M.ordered[,2] <- factor(M.ordered[,2], levels = as.character(1:15))
M.ordered[,3] <- as.numeric(as.character(M.ordered[,3]))
M.ordered[,4] <- as.numeric(as.character(M.ordered[,4]))
names(M.ordered) <- c("States", "States.num", "Prob", "Sum")

aux <- NULL
for(i in 1:dim(M.ordered)[1]){
  aux <- rbind(aux, vet.state[which(vet.state[,1] %in% as.character(M.ordered[i,1])),2:5])
}
names(aux) <- names(base)[3:6]

M.ordered <- cbind(M.ordered, aux)

setwd(path_results)
png(filename = "Pareto Plot for M.png", width = 12, height = 8, units = "in", res = 144)
ggplot() +
  geom_col(aes(x = as.numeric(as.character(M.ordered[,2])), y = M.ordered[,3])) +
  geom_line(aes(x = as.numeric(as.character(M.ordered[,2])), y = M.ordered[,4])) +

  annotate(geom = "text", x = rep(13, 16), y = 18:3,
    label = c(names(M.ordered)[1], as.character(M.ordered[,1]))) +
  annotate(geom = "text", x = rep(14, 16), y = 18:3,
    label = rep("-", 16)) +
  annotate(geom = "text", x = rep(15, 16), y = 18:3,
    label = c(names(M.ordered)[5], as.character(M.ordered[,5]))) +
  annotate(geom = "text", x = rep(17, 16), y = 18:3,
    label = c(names(M.ordered)[6], as.character(M.ordered[,6]))) +
  annotate(geom = "text", x = rep(19, 16), y = 18:3,
    label = c(names(M.ordered)[7], as.character(M.ordered[,7]))) +
  annotate(geom = "text", x = rep(21, 16), y = 18:3,
    label = c(names(M.ordered)[8], as.character(M.ordered[,8]))) +
  annotate(geom = "text", x = rep(22, 16), y = 18:3,
    label = rep("", 16)) +

  ylab("Probability [%]") +
  xlab("States") +
  scale_x_discrete(limits=as.character(M.ordered[,1]), rep("", 8))
dev.off()

#Graph for transition probability
png(filename = "HeatMap.png", width = 17, height = 8, units = "in", res = 144)
plot.heat(x = Mt, val = 30, val.most = 3)
dev.off()
setwd(path_base)

#J and correlation matrix
J <- base[,3:6]

#Kendall's correlation matrix
J.cor <- cor(x = J, method = "kendall")

```

```

#Cullen-Frey
setwd(path_results)
png(filename = "Cullen-Frey.png", width = 12, height = 12, units = "in", res = 144)
par(mfrow=c(2,2))
for(i in 1:dim(J)[2]){
  descdist.2(data = J[,i], boot = 200)
  title(names(J)[i])
}
dev.off()
par(mfrow = c(1,1))
setwd(path_base)

#Results for Cullen-Frey Graph
#Price = Logistic
#Energy = Beta

spot.price.fit <- fitdist(data = J[,1], distr = "logis", method = "mle")$estimate
future.price.fit <- fitdist(data = J[,2], distr = "logis", method = "mle")$estimate
spot.energy.fit <- fitdist(data = J[,3]/(max(J[,3])+0.001), distr = "beta", method =
"mle")$estimate
future.energy.fit <- fitdist(data = J[,4]/(max(J[,4])+0.001), distr = "beta", method =
"mle")$estimate

#save the fit dist
setwd(path_results)
capture.output(rbind(spot.energy.fit, future.energy.fit), file = "mat.fitdist.txt", append = FALSE)
capture.output(rbind(spot.price.fit, future.price.fit), file = "mat.fitdist.txt", append = TRUE)
setwd(path_base)

#Multivariate Copula
#copula class
myCop.copula <- archmCopula(family = "gumbel", dim = dim(J)[2])
#copula parameter
param.copula <- iTau(myCop.copula, mean(J.cor[lower.tri(J.cor)]))
#margins parameters
pmargins <- list(list(location = spot.price.fit[1], scale = spot.price.fit[2]),
  list(location = future.price.fit[1], scale = future.price.fit[2]),
  list(shape1 = spot.energy.fit[1], shape2 = spot.energy.fit[2]),
  list(shape1 = future.energy.fit[1], shape2 = future.energy.fit[2]))

#multivariate distribution class
myMvd1 <- mvdc(copula = archmCopula(family = "gumbel", param = param.copula, dim =
dim(J)[2]),
  margins = c("logis", "logis", "beta", "beta"),
  paramMargins = pmargins)

#generate copulas
n.copulas <- 1000 #number of copulas sample generated
var.copula <- rMvdc(n.copulas*dim(J)[1], myMvd1)
#percentile in copula
P.cop <- pMvdc(x = var.copula, mvdc = myMvd1)
var.copula <- na.omit(var.copula[(P.cop >= quantile(na.omit(P.cop), 0.975)),])

var.copula[,3] <- var.copula[,3] * (max(J[,3])+0.001)

```



```

var.copula[,4] <- var.copula[,4] * (max(J[,4])+0.001)

#remove negative numbers in copulas
var.copula[var.copula < 0] <- NA
var.copula <- as.data.frame(na.omit(var.copula))
names(var.copula) <- names(J)

#make state of generated copula
var.copula <- cbind(make.state(x = var.copula), var.copula)

names(var.copula)[1] <- "States"
J.copulas <- var.copula

#Remove states with prob = 0
J.copulas[,1] <- as.character(J.copulas[,1])
aux <- NULL
for(i in 1:dim(M)[1]){
  aux <- c(aux, which(J.copulas[,1] %in% as.character(M[i,1])))
}
J.copulas <- J.copulas[aux,]

#loop for states
#fit.hmm <- hmm(y = as.character(base[,2]), K=1)
#Number of Trials
n_trial <- 100
#trial structure
J.hmm.state <- NULL
#trial loop
for(i in 1:n_trial){
  J.hmm.state <- cbind(J.hmm.state, as.character(M[(sim.hmm(nsim = dim(J)[1], tpm =
as.matrix(Mt), Rho = M[,2])), 1]))
}

#Bootstrap Svar
#split the generated copulas
J.copulas.split <- split(J.copulas, J.copulas[,1])

#create structures for bootstraps coefficients
sp.coef <- NULL
fp.coef <- NULL
se.coef <- NULL
fe.coef <- NULL

#determine the optimal lag
lagmax <- 2

#loop in simulation
for(i in 1:dim(J.hmm.state)[2]){
  #loop in state
  J.sim <- data.frame(matrix(NA, nrow = dim(J)[1], ncol = dim(J)[2]))
  for(j in 1:length(J.copulas.split)){
    aux <- J.copulas.split[[j]]
    pos <- J.hmm.state[i] == aux[1,1]
  }
}

```

```

sample.pos <- sample(x = 1:dim(aux)[1], size = sum(pos), replace = TRUE)
J.sim[pos,] <- aux[sample.pos,2:5]
}
J.sim <- log(J.sim)
names(J.sim) <- names(J)

#VAR
J.var <- VAR(y = J.sim, p = lagmax, type = "none", exogen = cbind(Trend = base[,7], Trend2
= base[,8]))

#Spot price
sp.coef <- rbind(sp.coef, as.numeric(J.var$varresult$Spot.Price$coefficients))
#Future price
fp.coef <- rbind(fp.coef, as.numeric(J.var$varresult$Future.Price$coefficients))
#Spot energy
se.coef <- rbind(se.coef, as.numeric(J.var$varresult$Spot.Energy$coefficients))
#Future energy
fe.coef <- rbind(fe.coef, as.numeric(J.var$varresult$Future.Energy$coefficients))
}
sp.coef <- as.data.frame(sp.coef)
fp.coef <- as.data.frame(fp.coef)
se.coef <- as.data.frame(se.coef)
fe.coef <- as.data.frame(fe.coef)
names(sp.coef) <- names(J.var$varresult$Spot.Price$coefficients)
names(fp.coef) <- names(J.var$varresult$Future.Price$coefficients)
names(se.coef) <- names(J.var$varresult$Spot.Energy$coefficients)
names(fe.coef) <- names(J.var$varresult$Future.Energy$coefficients)

#Coefficient plots
#plot coef
setwd(path_results)

sp.gr <- plot.coef(sp.coef, ylb = "Spot Price Coefficients")
fp.gr <- plot.coef(fp.coef, ylb = "Future Price Coefficients")
se.gr <- plot.coef(se.coef, ylb = "Spot Energy Coefficients")
fe.gr <- plot.coef(fe.coef, ylb = "Future Energy Coefficients")

png(filename = "Coefficients.png", width = 18, height = 12, units = "in", res = 144)
grid.newpage()
pushViewport(viewport(layout = grid.layout(4,1)))
print(sp.gr, vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
print(fp.gr, vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
print(se.gr, vp = viewport(layout.pos.row = 3, layout.pos.col = 1))
print(fe.gr, vp = viewport(layout.pos.row = 4, layout.pos.col = 1))
dev.off()

#Multinomial Logistic Regression
set.seed(1) #seed
n_eff <- 1 #number of cases tested
n_boot <- n_trial #Number of bootstrap

#simulated states
y <- J.hmm.state
#most prob states
y.most <- (as.character(M.ordered[,1]))

```

```

#contextual variables
x <- base[,7:8]

#BOOTSTRAP multinomial logistic regression
#structures
temp_coef <- NULL
temp_pred <- NULL

#formula for regression
fml <- paste("State ~ ", names(x)[1], sep = "")
for(i in 2:dim(x)[2]){
  fml <- paste(fml, " + ", names(x)[i], sep = "")
}
fml <- as.formula(fml)

j<-1
while(j<=n_boot){
  TESTE <- try({
    co <- 1
    while(co == 1){
      ystar<- y[,j]
      C <- cbind(ystar, x)
      #only the most prob states
      C <- C[which(ystar %in% y.most),]
      C[,1] <- as.character(C[,1])

      names(C) <- c("State", names(x))
      Bfit<-multinom(fml, data = C, maxit = 1000)
      co <- Bfit$convergence
    }
    #Estimate
    temp_coef <- rbind(temp_coef, cbind(rownames(summary(Bfit)$coefficients),
summary(Bfit)$coefficients))

    #Predict values
    re_bfit <- as.data.frame(Bfit$fitted.values)
    temp_pred <- rbind(temp_pred, re_bfit)
  }, silent = TRUE)
  ifelse(is(TESTE, "try-error"), j<-j,j<-j+1)
}
coef_mlr <- as.data.frame(temp_coef)
pred_mlr <- as.data.frame(temp_pred)

#boxplot
setwd(path_results)
png(filename = "Boxplot for MLR.png", width = 12, height = 5, units = "in", res = 144)
boxplot(temp_pred, xlab = "States", ylab = "Probability")
dev.off()

#coefficient
coef_mlr.split <- split(coef_mlr, coef_mlr[,1])

#contextual variables
names.context <- names(x)

```

```

setwd(path_results)
png(filename = "Coefficients for MLR.png", width = 10, height = 12, units = "in", res = 144)
grid.newpage()
pushViewport(viewport(layout = grid.layout(7,2)))
print(plot.coef.mlr(x = coef_mlr.split[[1]], context = names.context, name.title =
"Coefficients"),
      vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
for(i in 2:7){
  print(plot.coef.mlr(x = coef_mlr.split[[i]], context = names.context),
        vp = viewport(layout.pos.row = i, layout.pos.col = 1))
}
for(i in 8:length(coef_mlr.split)){
  print(plot.coef.mlr(x = coef_mlr.split[[i]], context = names.context),
        vp = viewport(layout.pos.row = (i-7), layout.pos.col = 2))
}
dev.off()

```

Acknowledgements Comments from the Editor and two anonymous reviewers are gratefully acknowledged.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adeyemi, O. I., & Hunt, L. C. (2007). Modelling OECD industrial energy demand: Asymmetric price responses and energy-saving technical change. *Energy Economics*, 29, 693–709.
- Amorim, F., Vasconcelos, J., Abreu, C., Silva, P. P., & Martins, V. (2013). How much room for a competitive electricity generation market in Portugal? *Renewable and Sustainable Energy Reviews*, 18, 103–118.
- Asche, F., Misund, B., & Sikveland, M. (2018). The relationship between spot and contract gas prices in Europe. *Energy Economics*, 38, 212–217.
- Bashir, M. A., Sheng, B., Dogan, B., Sarwar, S., & Shahzad, U. (2020). Export product diversification and energy efficiency: Empirical evidence from OECD countries. *Structural Change and Economic Dynamics*, 55, 232–243.
- Beckmann, J., Belke, A., & Czudaj, R. (2014). Regime-dependent adjustment in energy spot and futures markets. *Economic Modelling*, 40, 400–409.
- Bekiros, S. D., & Diks, C. G. H. (2008). The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. *Energy Economics*, 30, 2673–2685.
- Bian, Y., Lv, K., & Yu, A. (2017). China's regional energy and carbon dioxide emissions efficiency evaluation with the presence of recovery energy: An interval slacks-based measure approach. *Annals of Operations Research*, 255, 301–321.
- Bonte, W., Nielsen, S., Valitov, N., & Engelmeyer, T. (2015). Price elasticity of demand in the EPEX spot market for electricity—New empirical evidence. *Economics Letters*, 135, 5–8.

- Cai, Y., Shi, X., Shao, H., Wang, R., & Liao, S. (2018). Energy efficiency state identification in milling processes based on information reasoning and Hidden Markov Model. *Journal of Cleaner Production*, 193, 397–413.
- Canelas, E., Pinto-Varela, T., & Sawik, B. (2020). Electricity Portfolio optimization for large consumers: Iberian electricity market case study. *Energies*, 13(9), 2249.
- Capitán-Herráiz, A., & Rodríguez-Monroy, C. (2013). Analysis of the traded volume drivers of the Iberian power futures market. *International Journal of Electrical Power and Energy Systems*, 44, 431–440.
- Chiou-Wei, S., Linn, S. C., & Zhu, Z. (2014). The response of U.S. natural gas futures and spot prices to storage change surprises: Fundamental information and the effect of escalating physical gas production. *Journal of International Money and Finance*, 42, 156–173.
- Ciarreta, A., & Zarraga, A. (2015). Analysis of mean and volatility price transmissions in the MIBEL and EPEX electricity spot markets. *The Energy Journal*, 36, 41–60.
- Cullen, A. C., & Frey, H. C. (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs*. Plenum Press.
- Danisman, O., & Kocer, U. U. (2021). Hidden Markov models with binary dependence. *Physical A: Statistical Mechanics and Its Applications*, 567, 12568.
- Delignette-Muller, M. L., & Dutang, C. (2014). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, 64, 1–34.
- Demirer, R., Kutun, A. M., & Shen, F. (2012). The effect of ethanol listing on corn prices: Evidence from spot and futures markets. *Energy Economics*, 34, 1400–1406.
- Dogan, B., Driha, O. M., Lorente, D. B., & Shahzad, U. (2021). The mitigating effect of economic complexity and renewable energy on carbon emissions in developed countries. *Sustainable Development*, 29, 1–12.
- Ewing, B. T., Hammoudeh, S. M., & Thompson, M. A. (2006). Examining asymmetric behavior in US petroleum futures and spot prices. *The Energy Journal*, 27, 9–24.
- Farooq, M. U., Shahzad, U., Sarwar, S., & Li, Z. (2019). The impact of carbon emission and forest activities on health outcomes: Empirical evidence from China. *Environmental Science and Pollution Research*, 26, 12894–12906.
- Fatima, T., Shahzad, U., & Cui, L. (2021). Renewable and nonrenewable energy consumption, trade and CO2 emissions in high emitter countries: Does the income level matter? *Journal of Environmental Planning and Management*, 64(7), 1227–1251.
- Ghazouani, A., Xia, W., Jebli, M. B., & Shahzad, U. (2020). Exploring the role of carbon taxation policies on CO2 emissions: Contextual evidence from tax implementation and non-implementation European countries. *Sustainability*, 12(20), 8680.
- Gil-Alana, L. S., Martin-Valmayor, M., & Wanke, P. (2020). The relationship between energy consumption and prices. Evidence from futures and spot markets in Spain and Portugal. *Energy Strategy Reviews*, 31, 100522.
- Huisman, R., Mahieu, R., & Schlichter, F. (2009). Electricity portfolio management: Optimal peak/off-peak allocations. *Energy Economics*, 31, 169–174.
- Kaleta, M., & Toczyłowski, E. (2008). Restriction techniques for the unit-commitment problem with total procurement costs. *Energy Policy*, 36, 2439–2448.
- Kaufmann, R. K., & Ullman, B. (2009). Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices. *Energy Economics*, 31, 550–558.
- King, M., & Cuc, M. (1996). Price convergence in North American natural gas spot markets. *The Energy Journal*, 17, 17–42.
- Li, G., Liu, W., Wang, Z., & Liu, M. (2017). An empirical examination of energy consumption, behavioral intention, and situational factors: Evidence from Beijing. *Annals of Operations Research*, 255, 507–524.
- Lin, S. X., & Tamvakis, M. N. (2001). Spillover effects in energy futures markets. *Energy Economics*, 23, 43–56.
- Lv, W., Hong, X., & Fang, K. (2015). Chinese regional energy efficiency change and its determinants analysis: Malmquist index and Tobit model. *Annals of Operations Research*, 228, 9–22.
- Maslyuk, S., & Smyth, R. (2009). Cointegration between oil spot and future prices of the same and different grades in the presence of structural change. *Energy Policy*, 37, 1687–1693.
- Mishra, V., & Smyth, R. (2016). Are natural gas spot and futures prices predictable? *Economic Modelling*, 54, 178–186.
- Moose, I. A., & Al-Loughani, N. E. (1995). The effectiveness of arbitrage and speculation in the crude oil futures markets. *The Journal of Future Markets*, 15, 167–186.
- Nelsen, R. B. (1999). *An introduction to copulas*. Springer.

- Nicolau, M., & Palomba, G. (2015). Dynamic relationships between spot and future prices. The case of energy and gold commodities. *Resources Policy*, 45, 130–143.
- Omrani, H., Shafaat, K., & Alizadeh, A. (2019). Integrated data envelopment analysis and cooperative game for evaluating energy efficiency of transportation sector: A case of Iran. *Annals of Operations Research*, 274, 471–499.
- Park, H., Mjelde, J. W., & Bessler, D. A. (2006). Price dynamics among U.S. electricity spot markets. *Energy Economics*, 28, 81–101.
- Perez-Franco, I., Garcia-Garcia, A., & Maldonado-Briegas, J. J. (2020). Energy transition towards a greener and more competitive economy: The Iberian case. *Sustainability*, 12, 3343.
- Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, 27(4), 1–32.
- Pindyck, R. S. (2001). The dynamics of commodity spot and futures markets: A primer. *The Energy Journal*, 22, 1–29.
- Pollin, R.; Chakraborty, S.; and Garrett-Peltier, H. (2015). An egalitarian clean energy investment program for Spain. Political Economy Research Institute, Working Paper no. 390.
- Sadorsky, P. (2000). The empirical relationship between energy futures prices and exchange rates. *Energy Economics*, 22, 253–266.
- Sarwar, S., Shahzad, U., Chang, D., & Tang, B. (2019). Economic and non-economic sector reforms in carbon mitigation: Empirical evidence from Chinese provinces. *Structural Change and Economic Dynamics*, 49, 146–154.
- Schmidt, R., & Stadtmüller, U. (2006). Nonparametric estimation of tail dependence. *Scandinavian Journal of Statistics*, 33(2), 307–335. <https://doi.org/10.1111/j.1467-9469.2005.00483.x>
- Shadzahl, U. (2020). Environmental taxes, energy consumption, and environmental quality: Theoretical survey with policy implications. *Environmental Science and Pollution Research*, 27, 24848–24862.
- Shahzad, U., Dogan, B., Sinha, A., & Fareed, Z. (2021). Does Export product diversification help to reduce energy demand: Exploring the contextual evidences from the newly industrialized countries. *Energy*, 214, 118881.
- Shahzad, U., Ferraz, D., Dogan, B., & Rebelatto, D. A. D. N. (2020). Export product diversification and CO2 emissions: Contextual evidences from developing and developed economies. *Journal of Cleaner Production*, 276, 124146.
- Silva, D., Dias, M. F., & Madaleno, M. (2019). Iberian electricity market spot and futures prices: Comovement and lead-lag relationship analysis. *International Journal of Sustainable Energy Planning and Management*, 19, 59–68.
- Silvério, R., & Szklo, A. (2012). The effect of the financial sector on the evolution of oil prices: Analysis of the contribution of the futures market to the price discovery process in the WTI spot market. *Energy Economics*, 34, 1799–1808.
- Talbi, B., Jebli, M. B., Bashir, M. F., & Shahzad, U. (2020). Does economic progress and electricity price induce electricity demand: A new appraisal in context of Tunisia. *Journal of Public Affairs*. (forthcoming).
- Ullah, I., Ahmad, R., & Kim, D. (2018). A prediction mechanism of energy consumption in residential buildings using hidden Markov model. *Energies*, 11(2), 358.
- Wang, Y., & Pham, H. (2012). Modeling the dependent competing risks with multiple degradation processes and random shock using time-varying copulas. *IEEE Reliability Society*, 61(1), 13–22. <https://doi.org/10.1109/TR.2011.2170253>
- Weron, R., & Zator, M. (2014). Revisiting the relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics*, 44, 178–190.
- Worthington, A., Kay-Spratley, A., & Higgs, H. (2005). Transmission of prices and price volatility in Australian electricity spot markets: A multivariate GARCH analysis. *Energy Economics*, 27(2), 337–350.
- Yan, G., & Trück, S. (2020). A dynamic network analysis of spot electricity prices in the Australian national electricity market. *Energy Economics*, 92, 104972.
- Yan, J. (2007). Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software*, 21(4), 1–21. <https://doi.org/10.18637/jss.v021.i04>
- Yang, L., & Yang, T. (2015). Energy consumption and economic growth from perspective of spatial heterogeneity: Statistical analysis based on variable coefficient model. *Annals of Operations Research*, 228, 151–161.
- Zhang, J., & Kotani, K. (2012). The determinants of household energy demand in rural Beijing: Can environmentally friendly technologies be effective? *Energy Economics*, 34, 381–388.
- Zhang, Y., & Wang, Z. (2013). Investigating the price discovery and risk transfer functions in the crude oil and gasoline futures markets: Some empirical evidence. *Applied Energy*, 104, 220–228.

- Zhou, P., Wu, F., & Zhou, D. Q. (2017). Total-factor energy efficiency with congestion. *Annals of Operations Research*, 255, 241–256.
- Zucchini, W., & Guttorp, P. (1991). A hidden Markov model for space-time precipitation. *Water Resources Research*, 27, 1917–1923.
- Zucchini, W., MacDonald, I. L., & Langrock, R. (2016). *Hidden Markov models for time series: an introduction using R* (Vol. 150). Boca Raton: CRC Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.