How are Day-ahead Prices Informative for Predicting the Next Day's Consumption of Natural Gas? Evidence from France

Arthur Thomas, Olivier Massol, Benoît Sévi

How are day-ahead prices informative for predicting the next day's consumption of natural gas? Evidence from France *

Arthur Thomas†
IFP Énergies Nouvelles & Université de Nantes

Olivier Massol‡
IFP Énergies Nouvelles, IFP School & City, University of London

Benoît Sévi §
Université de Nantes

This version: April 2019

Abstract

The purpose of this paper is to investigate whether the next day's consumption of natural gas can be accurately forecast using a simple model that solely incorporates the information contained in day-ahead market data. Hence, unlike standard models that use a number of meteorological variables, we only consider two predictors: the price of natural gas and the spark ratio measuring the relative price of electricity to gas. We develop a suitable modeling approach that captures the essential features of daily gas consumption and in particular the nonlinearities resulting from power dispatching. We use the case of France as an application as this is, as far as is known, the very first attempt to model and predict the country's daily gas demand. Our results document the existence of a long-run relation between demand and spot prices and provide estimates of the own- and cross-price elasticities. We also provide evidence of the pivotal role of the spark ratio which is found to have an asymmetric and highly nonlinear impact on demand variations. Lastly, we show that our simple model is sufficient to generate predictions that are considerably more accurate than the forecasts published by infrastructure operators.

JEL Classification: L95; Q41; Q47; C22; C53
Keywords: Natural gas markets, day-ahead prices, demand price elasticity, load forecasting

* We are indebted to Kevin Forbes and Yannick Le Pen for insightful discussions and suggestions. We have also benefited from helpful discussions with conference participants at ISEFI (Paris, 2018) and TAI (Washington DC, 2018). Of course, any remaining errors are ours. This research has been supported by the Chair "The Economics of Natural Gas." The views expressed herein are strictly those of the authors and should not be construed to reflect the opinions of IFP Énergies Nouvelles.

† Department of Economics, IFP Énergies Nouvelles, 1-4 av. de Bois Préau, F-92852 Rueil-Malmaison, & LEMNA, Université de Nantes, BP. 52231 Chemin de la Censive du Tertre, 44322 Nantes Cedex, France E-mail: arthur.thomas@ifpen.fr

‡ IFP Énergies Nouvelles, 1-4 av. de Bois Préau, F-92852 Rueil-Malmaison, & Center for Energy Economics and Management, IFP School, 228-232 av. Napoléon Bonaparte, F-92852 Rueil-Malmaison, France, & Department of Economics, City, University of London, Northampton Square, London EC1V 0HB, UK.

§ LEMNA, Université de Nantes, BP. 52231 Chemin de la Censive du Tertre, 44322 Nantes Cedex, France.

1 Introduction

The accuracy of the gas demand forecasts issued by Transmission System Operator (TSO) is now becoming an important issue in regulatory debates and has motivated the adoption of dedicated incentive schemes in some countries.[1] In response, TSOs have implemented advanced forecasting tools combining several methodologies (e.g., time series, neural network, adaptative logic networks) along with a plethora of variables (e.g., temperatures, wind speeds, rain, snow, cloud cover, forecasted power demand). Yet, despite these efforts, forecasting the next day's consumption of natural gas remains a challenging task.[2] By testing an alternative forecasting approach based on the informational content of day-ahead prices, the present paper usefully contributes to the ongoing discussion on the performance of short-term consumption forecasts used in the gas industry.

Over the last two decades, a series of European regulatory reforms have prompted the emergence of a collection of day-ahead wholesale markets for natural gas, the so-called "gas hubs," that turned out to become an important source of gas procurement as the previously monopolized industry structure gradually became more fragmented (Miriello and Polo, 2015). By construction, these markets have been developed to cope with local network balancing needs and allow an optimal scheduling of resources. Their functioning is thus closely affected by the detailed balancing rules used by the TSO. An important milestone in the design of these balancing procedures occurred in 2014 when the European Commission imposed a unified network code on TSOs.[3] Yet, despite that harmonization, market analysts recurrently point significant differences in the perceived degree of trading liquidity observed at the European gas hubs (Heather and Petrovich, 2017). A fundamental public policy issue is, thus, whether the current market design generates transparent spot prices that reflect the market participation of all concerned economic agents (suppliers, trading firms, and consumers).

In the electricity sector, Forbes and Zampelli (2014) proposed an original approach to examine the infor-

[1] For example, in the UK, a dedicated annual incentive scheme has been implemented so that, depending on the observed average annual forecast error, the TSO can earn up to a maximum of £10 million (in case of 100% accuracy of the published day-ahead demand forecasts) or lose up to £1.5 million (National Grid, 2018). In Italy, the regulatory authority monitors the forecasting error of the next day's load and uses it as a performance indicator to assess the quality of the information transmitted to the market (ENTSOG, 2017 p.36).

[2] For example, in the French southern gas balancing zone, the root mean squared error of the day-ahead consumption forecasts issued at 5pm by GRTGaz – the largest TSO – was approximately 23 GWh over the year 2015 corresponding to about 6.6% of the average daily load. That year, one working day out of four (the exact proportion is 25.2%) experienced an absolute forecasting error larger than 5% in relative terms (source: smart.grtgaz.com).

[3] See Commission Regulation (EU) No 312/2014 of March 26, 2014, OJ L 91, 27.3.2014.

mational content of day-ahead electricity prices. They hypothesize that if day-ahead markets for electricity are efficient then these prices should reflect the processed information of all market participants regarding the next day's load. That consideration led them to test whether it was possible to improve the predictions of the next day's electricity load using solely the information contained in the day-ahead price series. As an application, they examine California's PG&E aggregation area and apply traditional time-series techniques (namely a linear ARMAX specification) to model the next day's load as a function of a single explanatory variable: the day-ahead spark ratio defined as the electricity to gas price ratio. Their results reveal that this approach is sufficient for computing very accurate forecasts which outperform those published by the system operators. Remarkably, their results document the highly significant informational content of day-ahead prices in the case of electricity.

Our paper is, to the best of our knowledge, the first econometric study of the daily interactions between day-ahead prices and the natural gas demand observed at a given hub. In some respects, it extends Forbes and Zampelli's (2014) approach in highlighting what necessary specific dimensions must be considered to produce accurate natural gas demand forecasts. Noticeably, we provide evidence that this form of energy necessitates an adapted modeling strategy which differs from the one used for electricity. Essentially, three main characteristics are typical of the gas market: (i) the fact that the aggregate gas demand emanates from both end-users and thermoelectric generation, (ii) the expected nonlinearity in the relationship between price and demand incidental to the level of the relative price of electricity to natural gas (spark ratio), and (iii) the unit-root properties of the data. We further elaborate on these three distinctive features in section 2 as they provide the essential justifications for our modeling choice and are thus key in our analysis. We simply highlight here that, in light of these features, we consider two nonlinear specifications that are extensions of the well-known Autoregressive Distributed Lag (ARDL) model: the Nonlinear ARDL (NARDL) and a new one, that we propose here, the Threshold ARDL (TARDL). By doing so, we investigate the presence of a long-run relationship between consumption, day-ahead price, and spark ratio and explore the potential asymmetric influence of the spark ratio on observed consumption levels.[4]

As an application, we examine the case of the two French wholesale markets – namely, the Point d'Échange

[4]Notably, our modeling approach does not include any meteorological variables as it posits that all available meteorological information should be reflected in the trading decisions taken by market participants and thus in the day-ahead prices. It should be emphasized that, following Forbes and Zampelli (2014), our aim is to explore the informational content of day-ahead markets rather than investigating the potential contribution of additional variables to the prediction of gas consumption.

de Gaz Nord (PEG Nord) and the Trading Region South (TRS) – over the 2015–2018 period. This allows us to present a series of original findings. First, we provide evidence of a symmetric and significant long-run relationship between the daily demand level, the spot price of natural gas, and the relative price of electricity to gas. Second, we document the magnitude of the demand price elasticities. Third, we show that, in the short run, the spark ratio has an asymmetric and nonlinear impact on observed demand levels. In each market, the reported relationship obtained with the TARDL model is sufficiently robust to produce day-ahead forecasts that are considerably more accurate than that published by network operators.

Our main contribution to the literature is to show that public information, such as day-ahead prices, can be used to produce efficient forecasts of tomorrow's consumption by relying on an uncomplicated econometric model. The fact that our demand forecasts are much better than those provided by TSOs appears quite puzzling as TSOs are expected to hold superior information. Why such superior information does not translate into better demand predictions remains an open question.

Our framework can provide useful guidance to a large audience interested in the dynamics of natural gas demand in the short run and in the reaction of that demand to market prices. While a large applied econometrics literature has approached the question using medium to low frequency data (e.g., monthly, quarterly or annual), that reaction has never been examined using daily data. In principle, the use of daily data provides a much richer data set for eliciting short-run effects from lagged changes in prices on the observed demand. Geweke (1978) stresses that estimation over broader data intervals can result in significant bias. His analysis indicates that aggregation over time can create a type of omitted variables bias problem because the intertemporal lag distribution is not properly specified. In our case, the use of daily data may provide more reliable estimates of the price elasticity of natural gas demand, particularly if natural gas demand responds rapidly to energy price changes. As that elasticity parameter plays a very important role in the models developed to examine the impacts of a possible sudden temporary disruption in gas supplies on optimal import policies,[5] our modeling approach usefully contributes to the policy discussions related to the security of foreign-controlled gas supplies in importing nations.

Though our discussion is confined to the French case, we believe that the results are pertinent for other countries engaged in a transition toward less carbon-intensive energy systems. In France, the gas consump-

[5]See: Manne et al. (1986), Hoel and Strøm (1987), Markandya and Pemberton (2010), and Abada and Massol (2011).

tion emanating from the power sector exhibits large and sudden variations because Combined Cycle Gas Turbines (CCGT) plants are primarily dispatched as peaking units, which leads to large flow variations in the gas network as these plants ramp up and down. That situation is likely to prefigure the new role assigned to gas-fired power plants when a previously thermoelectric dominated power system experiences a massive penetration of renewable generation. Because of their almost zero marginal costs of production, solar and wind generators are placed at the beginning of the electricity 'merit order,' which greatly reduces the need to dispatch gas-fueled generators as baseload or mid-merit units (Green and Vasilakos, 2010) and thus leads to large variations in the gas demand emanating from these plants (Qadrdan et al., 2010).

The present analysis has important implications for the cost-efficient operation of a natural gas pipeline system and its economic regulation in the broader context of decarbonized energy systems. As gas-fired generation is increasingly used as a backup technology in power systems, the short-term variability observed in the power sector is increasingly transferred to gas infrastructures. Because of both the increased variability of the next days' loads and the possibility of demand forecasting errors, gas TSOs' are forced to adopt precautionary network management strategies based on the buildup and discharge of a pipeline inventory named linepack (Gopalakrishnan and Biegler, 2013; Tran et al., 2018). The linepack storage is an important source of short-term flexibility that could be extremely valuable in decarbonized energy systems as shown in Sun et al. (2012) and Arvesen et al. (2013). Yet, from a market design perspective, that storage service is seldom sold to market participants in Europe and a substantial share of its cost is socialized by means of transport tariffs (Keyaerts et al., 2011; Hallack and Vazquez, 2013). Because of these factors, there is a heightened interest in the accuracy of the gas demand forecasts published by the TSOs. That topic is now considered to be an important regulatory policy issue that has recently motivated the implementation of specific incentive schemes in some countries (e.g., the UK and Italy).

The remaining sections of this paper are organized as follows. In the next section, we present some attributes that are typical to gas markets and which inspire our econometric approach which is exposed in section 3. In section 4, we present the data along with some preliminary analysis of the series. The empirical findings are provided in section 5. Section 6 concludes.

2 Distinctive features in forecasting short-term natural gas prices

At least three distinctive features are worth highlighting when attempting to model and forecast the daily consumption of natural gas. The first one is related to the specific nature of natural gas as a source of energy. The original work of Forbes and Zampelli (2014) focuses on electricity, which is a type of energy predominantly consumed by end-users of that specific form of energy. In contrast, natural gas is either directly consumed by end-users of that specific form of energy (e.g., by households or industries to produce heat) or is converted into a secondary form of energy, namely electricity. While one can assume that the consumption emanating from the former users is directly influenced by the price of natural gas, those from the power sector are likely to be influenced by the relative price of natural gas and electricity. Given the importance of the power sector's demand for natural gas, one can hardly overlook the role of electricity prices in the observed aggregate demand of natural gas. Hence, the present analysis suggests considering two variables to predict the next day's load: the day-ahead price of natural gas and the day-ahead spark ratio. As in Forbes and Zampelli (2014), our results provide evidence of the key role of the spark ratio in predicting tomorrow's consumption.

The second feature is related to the use of natural gas in power generation. On a given power system, gas-fueled generation is seldom the unique technology available to generate electricity. Thus, depending on the observed level of the relative price of electricity and gas, it is likely that the consumption of natural gas in the power sector differs. Arguably, for low levels of the spark ratio, the revenue derived from gas-fueled generation is not sufficient to compensate both the thermal losses and other operating costs. One can thus conjecture that the consumption of natural gas in the power sector remains circumscribed to a few cogeneration plants (i.e., Combined Heat and Power plants) that must run to supply heat at industrial sites or district heating systems and is thus not very responsive to the electricity/gas ratio. In contrast, whenever the spark ratio is large enough, large gas-fueled thermal power stations consume natural gas, which suggests a stronger positive relationship between the gas load and the spark price ratio. Overall, that discussion suggests opting for a modeling approach that can incorporate non-linearities and is suitable for detecting the possible presence of asymmetries between the observed consumption levels of natural gas and the value of the spark ratio. In our application, our estimates strongly support the nonlinear assumption. Indeed, nonlinearity is shown to be critical to our analysis and estimates related to nonlinearity are consistently highly significant.

The third specific feature has a rather methodological nature. In recent years, a large empirical literature has examined the time series properties of day-ahead electricity prices.[6] In these studies, spot power prices are commonly found to be mean reverting and unlikely to have a unit root. In contrast, the day-ahead price series of natural gas (see De Vany and Walls, 1993; Serletis and Herbert, 1999; Renou-Maissant, 2012; or Thoenes, 2014, among others) and the daily consumption of natural gas, as examined in Giulietti et al. (2012), are often found to be non-stationary and an integrated process of order one, $I(1)$. As a result, one cannot directly estimate a regression equation involving the variables in levels without conveying the risk of generating spurious results. To overcome that problem, we consider the ARDL model suggested in Pesaran and Shin (1998) and Pesaran et al. (2001) which is a single cointegration and error correction approach that yields valid results regardless of whether the underlying variables are integrated in different orders (one, zero, or a combination of both). This method has two main advantages. First, it allows us to test for the presence of a long-run relationship among the variables without any prior knowledge on their order of integration, which partly avoids problems associated with unit root testing.[7] Second, it offers a parsimonious modeling approach that can easily be extended to incorporate nonlinearities using partial sum decompositions as in the Nonlinear ARDL (NARDL) model proposed by Shin et al. (2014) or the Threshold ARDL (TARDL) model developed in this paper. We provide evidence that our data series either have a unit root or are stationary, thereby justifying the use of an ARDL specification.

Overall, as will be shown below, our empirical results confirm that accounting for these three features is very important for modeling and accurately predicting the next day's consumption of natural gas. In this respect, our empirical modeling approach thus noticeably departs from the one in Forbes and Zampelli (2014) although our aim remains quite similar in spirit.

[6] See, e.g., Serletis and Herbert (1999), Lucia and Schwartz (2002), Knittel and Roberts (2005), Worthington et al. (2005), Bunn and Gianfreda (2010), de Menezes and Houllier (2016), Gianfreda and Bunn (2018).

[7] As will be discussed below, unit-root testing is nevertheless useful for checking that the series are not $I(2)$, i.e., integrated or order 2.

3 Econometric approach

In this section, we first provide a condensed review of the standard ARDL model before presenting two extensions: the NARDL and the TARDL. We let q_t denote the quantity demanded on day t in a given wholesale market.[8] We aim to model the quantity demanded as a function of: the day-ahead price of natural gas that is delivered that day p_t, the spark ratio $s_t := p_t^E/p_t$ where p_t^E is the day-ahead price of electricity delivered during the peakload block of day t, and Winter_t is a deterministic seasonal dummy variable that takes the value one during the gas winter and zero otherwise. Hence, our model has the following form: $q_t = f\left(p_t, s_t, \text{Winter}_t\right)$.

3.1 ARDL model

The linear ARDL model of Pesaran et al. (2001) enables interpretation based on the short- and long-run effects of the explanatory variables on the dependent variable. This approach has an important advantage over other cointegration techniques such as the ones of Engle and Granger (1987) or Johansen and Juselius (1990), as it can be applied regardless of whether all the variables share the same order of integration, which is not possible under alternative cointegration models.[9]

In our framework, the specification of a linear ARDL model is as follows:

$$\begin{aligned}\Delta q_t = {} & \alpha + \rho q_{t-1} + \theta_1 s_{t-1} + \theta_2 p_{t-1} + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} + \sum_{i=0}^{q_1-1} \gamma_i \Delta s_{t-i} \\ & + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i} + \kappa \text{Winter}_t + \zeta \Delta \text{Winter}_t + \mu_t\end{aligned} \tag{1}$$

where: Δ is the first difference operator; α denotes an intercept; ρ is the feedback coefficient (expected to be negative); θ_1 and θ_2 represent the long-run coefficients; ϕ_i, γ_i and δ_i are the short-run coefficients; p, q_1 and q_2 are the respective lag orders for the dependent and explanatory variables; κ and ζ represent the long

[8] For simplicity, we abstract from indexing variables with respect to the market under scrutiny. Our notation system works in the same way for the two markets examined in the application presented below.

[9] In an ARDL model, the variables can be either stationary (i.e., integrated of order zero $I(0)$) or integrated of order one $I(1)$. However, this model is not valid when there are $I(2)$ variables.

and short-run effect of the gas winter variable; and μ_t is the error term. These coefficients can be combined to obtain the long-run multipliers[10] $\beta_1 := -\theta_1/\rho$ and $\beta_2 := -\theta_2/\rho$. The coefficients γ_i (respectively δ_i) capture the short-run adjustments of natural gas demand to spark ratio (respectively gas price) shocks. In particular, γ_0 and δ_0 measure the contemporaneous impacts of these changes on natural gas demand variations.

The ARDL model can be used to examine the existence of a long-run (i.e., cointegration) relationship among the underlying variables by employing the bound testing approach described in Pesaran et al. (2001). This amounts to testing the null hypothesis of no cointegration among the underlying variables, that is, $H_0 : (\rho = \theta_1 = \theta_2 = 0)$. To test this hypothesis, the authors propose a non-standard F-test that takes into account the stationarity properties of the variables and evaluate bounds for the critical values at any significance level. The lower bound assumes that all variables are $I(0)$, whereas the upper bound assumes that all variables are $I(1)$. If the test statistic is larger than the upper bound critical value, the null hypothesis of no cointegration is rejected, which means that a cointegrating relationship among the underlying variables can be ascertained. Conversely, if the test statistic is lower than the lower bound then this null hypothesis is not rejected, which indicates that the underlying variables are not cointegrated. Lastly, if the test statistic falls between these two bounds, the inference remains inconclusive.

When the variables are in natural logarithms, as in the application discussed below, a straightforward economic interpretation can be given to the long-run coefficients. By construction, the long-run elasticity of natural gas demand to electricity price is β_1, which is expected to be positive, as one can expect both energies to be substitutable. Also, the long-run price elasticity of natural gas demand is $\beta_2 - \beta_1$, which is expected to be negative as one can conjecture that natural gas is a normal good.

3.2 NARDL model

Although the ARDL model enables the investigation of the short- and long-run relationships between variables, it presumes that all exogenous variables have symmetric effects on the dependent variable and thus becomes unsuitable when these linkages are nonlinear and/or asymmetric. To overcome that limitation,

[10]By gathering the terms in levels, one can use that definition of the multipliers to express the specification in the usual ECM form: $\Delta q_t = \alpha + \rho\,(q_{t-1} - \beta_1 s_{t-1} - \beta_2 p_{t-1} - \beta_3 \text{Winter}_t) + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} + \sum_{i=0}^{q_1-1} \gamma_i \Delta s_{t-i} + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i} + \zeta \Delta \text{Winter}_t + \mu_t$, where $\beta_3 := -\kappa/\rho$ is the long-run multiplier associated with the winter dummy variable (expected to be positive).

Shin et al. (2014) developed the nonlinear autoregressive distributed lag (NARDL) model that offers an asymmetric expansion of the original ARDL model.

In the NARDL model, short-run and long-run nonlinearities are introduced via positive and negative partial sum decompositions of the explanatory variables. In the present paper, we focus solely on the asymmetric impact that the spark ratio s_t can have on the consumption of natural gas.[11] We thus decompose the spark ratio s_t as follows:

$$s_t = s_0 + s_t^+ + s_t^- \tag{2}$$

where s_0 is an arbitrary initial value and s_t^+ and s_t^- denote partial sum processes which accumulate positive and negative changes in s_t respectively. These partial sum processes are derived from the first differences $\Delta s_j := s_{j+1} - s_j$ using the following definitions:

$$s_t^+ = \sum_{j=1}^{t} s_j^+ \quad \text{with} \quad \forall j, \quad s_j^+ := max\left(\Delta s_j, 0\right)$$

$$s_t^- = \sum_{j=1}^{t} s_j^- \quad \text{with} \quad \forall j, \quad s_j^- := min\left(\Delta s_j, 0\right)$$

This decomposition is then introduced in (1) to obtain the NARDL model, which is an immediate generalization of the genuine ARDL but allows for the presence of short- and long-run asymmetries which are seen to be very useful in our case. The NARDL model can be written as follows:

$$\begin{aligned}\Delta q_t = \alpha + \rho q_{t-1} + \left(\theta_1^+ s_{t-1}^+ + \theta_1^- s_{t-1}^-\right) + \theta_2 p_{t-1}^G + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} \\ + \sum_{i=0}^{q_1-1} \left(\gamma_i^+ \Delta s_{t-i}^+ + \gamma_i^- \Delta s_{t-i}^-\right) + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i}^G + \kappa Winter_t + \zeta \Delta Winter_t + \mu_t\end{aligned} \tag{3}$$

where θ_1^+ and θ_1^- are the long-run asymmetric coefficients and γ_i^+ and γ_i^- are the short-run asymmetric co-

[11] Our focus on the asymmetric response of gas consumption to the spark-ratio is derived from a series of empirical investigations conducted on the two French wholesale markets. In both cases, we consistently observed that the null hypothesis of a symmetric response of gas consumption to variations in the day-ahead price of natural gas was always rejected in the short run, but not in the long run. These results, which motivate our econometric specifications, are available from the authors upon request.

efficients representing the contemporaneous impacts of positive and negative changes in the spark ratio on natural gas demand variations. The asymmetric long-run multipliers are $\beta_1^+ := -\theta_1^+/\rho$ and $\beta_1^- := -\theta_1^-/\rho$.

In this framework, the non-standard bounds-based F-test of Pesaran et al. (2001) are still valid for examining the presence of an asymmetric long-run relationship among the variables in levels. More specifically, the null hypothesis becomes $H_0 : \big(\rho = \theta_1^+ = \theta_1^- = \theta_2 = 0\big)$.

By construction, the NARDL specification nests three special cases: (i) a symmetric long-run relationship, that is, the null hypothesis $H_0 : \big(\theta_1^+ = \theta_1^-\big)$; (ii) a symmetric short-run relationship, that is: $H_0 : \big(\gamma_i^+ = \gamma_i^-, \forall i \in \{1, ..., (q_1 - 1)\}\big)$; and (iii) the joint presence of long- and short-run symmetry as in the original ARDL model. These three restrictions can be tested using standard Wald tests.

3.3 Threshold ARDL model

The NARDL model implicitly uses a zero threshold value to define the partial sum processes as an observed variation is thought to be either positive or negative. While the use of such a zero threshold value may be appealing in macroeconomics or in finance,[12] one can question its relevance for the present application. Arguably, the technological considerations in section 2 suggest that the observed level of the spark ratio is likely to have a nonlinear influence on the power sector's demand for natural gas. Indeed, for low levels of that ratio, the impact is likely to be of minor importance. The other way round, when the value of the spark ratio is large enough to compensate the costs for generating electricity at large gas-fired power stations, we expect it to have a larger influence on gas demand.

These considerations lead us to consider a different decomposition of the exogenous variable s_t that explicitly refers to a possibly non-zero threshold. Several recent contributions in economics have suggested decomposing the explanatory variable to allow for likely different regimes.[13] By construction, these earlier approaches compare the value of the first-differenced variable (here Δs_t) with a threshold and directly use that comparison to define the partial sum processes. However, the discussion above suggests that in the

[12] For instance, to capture the potential asymmetries occurring during expansionary or contractionary periods of the business cycle or to model the effect of positive and negative financial news

[13] See, e.g., Greenwood-Nimmo et al. (2011), Pal and Mitra (2015), Bagnai et al. (2018).

present case, it would be preferable to examine whether an observed variation in the spark ratio has or not the same impact on the observed demand variation if the level attained by that explanatory variable exceeds or not a given threshold Th. Formally, this leads us to consider the following new decomposition of the spark ratio s_t:

$$s_t = s_0' + s_t^{>Th} + s_t^{\leq Th} \tag{4}$$

$$s_t^{>Th} = \sum_{j=1}^{t} \Delta s_j^{>Th} = \sum_{j=1}^{t} \Delta s_j \mathbb{I}_{s_j > Th} \tag{5}$$

$$s_t^{\leq Th} = \sum_{j=1}^{t} \Delta s_j^{\leq Th} = \sum_{j=1}^{t} \Delta s_j \left(1 - \mathbb{I}_{s_j > Th}\right) \tag{6}$$

where s_0' is an arbitrary initial value (different from the one obtained in the NARDL) and $\mathbb{I}_{s_j > Th}$ is the indicator function that takes the value 1 if the condition $s_j > Th$ is satisfied and 0 otherwise.

The specification of the Threshold ARDL model is then obtained by introducing the decomposition in (1), that is:

$$\begin{aligned} \Delta q_t = \alpha + \rho q_{t-1} + \left(\theta_1^{>Th} s_{t-1}^{>Th} + \theta_1^{\leq Th} s_{t-1}^{\leq Th}\right) + \theta_2 p_{t-1}^G + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} \\ + \sum_{i=0}^{q_1-1} \left(\gamma_i^{>Th} \Delta s_{t-i}^{>Th} + \gamma_i^{\leq Th} \Delta s_{t-i}^{\leq Th}\right) + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i}^G + \kappa Winter_t + \zeta \Delta Winter_t + \mu_t \end{aligned} \tag{7}$$

where $\theta_1^{>Th}$ and $\theta_1^{\leq Th}$ are the long-run asymmetric coefficients and $\gamma_i^{>Th}$ and $\gamma_i^{\leq Th}$ are the short-run asymmetric coefficients. The associated asymmetric long-run multipliers are $\beta_1^{>Th} := -\theta_1^{>Th}/\rho$ and $\beta_1^{\leq Th} := -\theta_1^{\leq Th}/\rho$.

For a given value of the threshold parameter Th, that specification can be estimated using OLS. The value of that threshold parameter can be determined using a grid search algorithm over the spark-ratio variable. In this paper, we consider the grid set defined by the percentiles of s_t – after trimming for the 10^{th} and 90^{th}

percentiles so as to maintain a sufficient number of observations in each regime – and select the threshold value that minimizes the sum of squared residuals.

Here again, the long-run and short-run symmetry restrictions – hence the null hypotheses $H_0 : \left(\theta_1^{>Th} = \theta_1^{\leq Th}\right)$ and $H_0 : \left(\gamma_i^{>Th} = \gamma_i^{\leq Th}, \forall i \in \{1, ..., (q_1 - 1)\}\right)$ – can be tested using the Wald statistic.

4 Preliminary Analysis of the Data

4.1 Data

We focus on the two gas balancing zones in France, namely PEG Nord and TRS, that are respectively associated with the country's northern and southern wholesale markets for natural gas. In France, transit to and from other countries represents a modest share of the flows transported on the gas pipeline systems. One can thus expect that the price formed at such a hub reflects the interaction of supply and demand prevailing in that specific balancing zone. From a regulatory perspective, these two gas hubs share a common market design and are monitored by the same regulatory authority: the Commission de Régulation de l'Énergie (CRE). Despite that institutional similarity, PEG Nord attracts a greater number of active market participants and is reputed to provide a more liquid trading environment than TRS (Heather and Petrovich, 2017).

We consider the period covering April 1, 2015, to September 30, 2018. The starting date is the day at which trading operations commenced at the TRS balancing zone (CRE, 2012). During that period, the PEG Nord and the TRS experienced a steady institutional environment comprising unchanged infrastructure access rules and balancing conditions. In both markets, the main provisions stipulated in the EU's network code on gas balancing of transmission networks were already implemented at that time (ACER-ENTSOG, 2014).

The daily data used in this paper are collected from the sources indicated in Table 1. GRTGaz is the unique TSO operating in the PEG Nord balancing zone and provides the consumption data observed in that zone. In contrast, two TSOs operate in the TRS region – GRTGaz and Teréga (formerly named TIGF) – and our daily consumption series is obtained by adding the figures reported by these two TSOs. In France, there is a unique wholesale market for electricity that covers the entire metropolitan territory. We use the day-ahead price of electricity for the peakload block that covers the hours from 9am to 8pm on the next working day. In each gas balancing zone, the spark ratio is computed by dividing the day-ahead electricity price by the corresponding day-ahead price of natural gas.

Following the approach retained in numerous empirical analyses of energy markets (e.g., Ramanathan et al.,1997; Karakatsani and Bunn, 2008; Bordignon et al., 2013), we concentrate our attention on the working days and remove the weekends from the data.[14] Hence, the day-ahead gas prices formed on a Friday value gas to be delivered on the following Monday. Arguably, the informational content of such a day-ahead price could be lower than that formed on the other days for gas to be delivered during the next day. In all specifications, we thus supplement the constant parameter by four daily dummy variables in order to control for possible day-of-the-week effects.

The whole dataset has a total of 640 observations. In the sequel, that data set is further divided into two parts. The first part, covering the period April 1, 2015,-December 31, 2016 (i.e, 406 observations), is only used for model estimation. The remaining part comprises 234 observations and is used for evaluating out-of-sample forecasts.

Table 1
Data sources.

Market	**Data**	**Source**	**Specification**	**Unit**
Natural gas PEG	Daily consumption	smart.grtgaz.com (GRTGaz)	Consumption North	GWh
	Day ahead price	Bloomberg	PEG Nord, End of Day	€/MWh
Natural gas TRS	Daily consumption	smart.grtgaz.com (GRTGaz) opendata.reseaux-energies.fr (Terega)	Consumption South	GWh
	Day ahead price	Bloomberg	TRS, End of Day	€/MWh
Electricity	Day ahead price	Bloomberg	Powernext Peakload, End of Day	€/MWh

[14]Moreover, there is only one unique price quote for the entire weekend (i.e., two days).

4.2 Descriptive statistics

Figure 1 provides plots of the price and consumption series in levels for the entire sample period. A visual inspection of the consumption plot indicates that the daily consumption observed in PEG Nord is consistently larger than that measured in the TRS balancing zone, which is not surprising as it is home to a larger share of the population, gathers the biggest industrial sites, and is generally affected by cooler weather. In both markets, the series share a similar pattern with peaks during the winter, when cold weather increases the demand for natural gas heating.

In the natural gas industry, the period between November 1 and March 31 is officially designated as the "gas winter."[15] That period is representative of the heating season and plays a specific role in the regulatory framework governing the operations of the natural gas industry. In France, these dates are mentioned in the seasonal storage obligations imposed on both gas retailers and TSOs[16] and are also explicitly used by gas storage operators to define their seasonal storage services. For example, the commercial seasonal storage services sold by Storengy – the country's largest operator of underground storage sites – are such that the inventory level attained on November 1 cannot be lower than 85% of the storage volume purchased and that the level measured on March 31 cannot exceed 40% of that volume (source: storengy.com). To control for that seasonal effect, we define the step dummy variable Winter_t that takes the value one during the gas winter and zero otherwise.

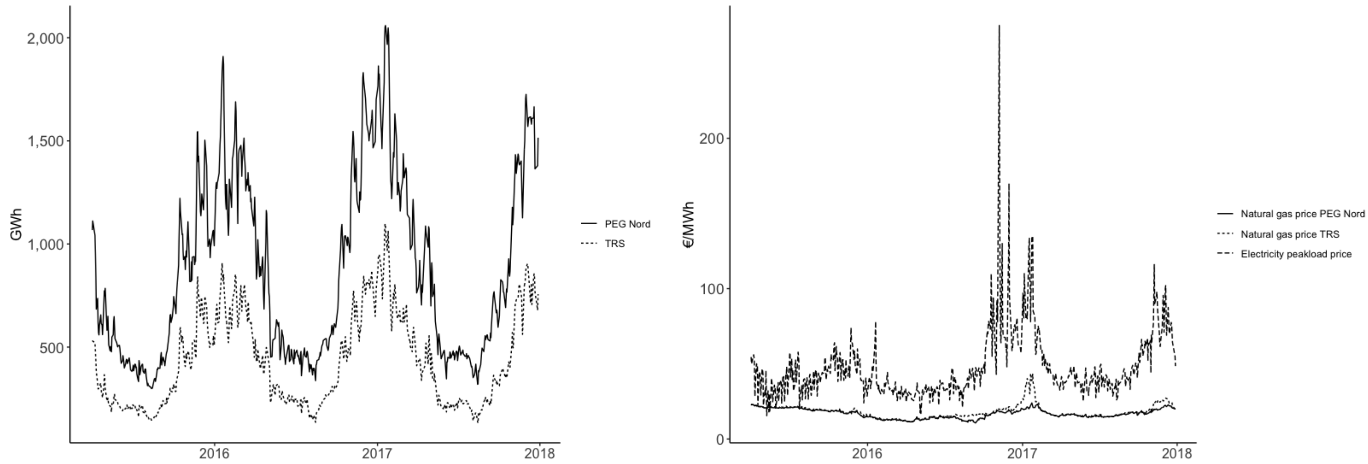
The price plots show that the day-ahead price of electricity measured during the peakload block is both larger and more variable than that of natural gas. One can also note that the prices formed at the southern gas hub (TRS) are very similar to the ones at PEG Nord (except during the winter 2017–2018).

Table 10 (in Appendix A) summarizes the descriptive statistics for the series in levels during the estimation period. The average consumption figures indicate that the size of the PEG Nord zone is roughly twice larger than that of TRS (i.e., 814.8 GWh compared with 401.8 GWh). The coefficient of variation of these two series are quite large and attain 50.6% and 52.1% in PEG Nord and TRS, respectively. The slightly larger

[15] That definition is not specific to France. For example, the discussion in Uría and Williams (2007) shows that these dates are also used to delineate the official storage period in California's natural gas industry.

[16] In France, the government monitors the aggregate inventory level at the domestic underground storage sites and explicitly imposes the industry to attain a predefined level on November 1 to preserve the security of supply during the next winter. It should be noted that this regulatory obligation is not specific to France as Uría and Williams (2007) indicate that a similar mechanism also exists in California.

Figure 1
Daily consumption of natural gas (left) and day-ahead energy prices in France (right).



coefficient observed in the southern gas balancing zone could be related to that region's smaller industrial base, as the consumption observed in large industrial sites usually exhibits limited seasonal variations.

The distributional properties of these series show some signs of non-normality. All the series are positively skewed. Moreover, the electricity price series displays significant excess kurtosis, which means that it has fatter tails than a normal distribution (Knittel and Roberts, 2005). For all series, the presence of non-normal distributions is checked through the highly significant Jarque-Bera statistics. Thus, in what follows, all the variables are transformed into natural logarithms to facilitate the economic interpretation of the estimated coefficients and to partly deal with the fact that the empirical distributions are not Gaussian.

4.3 Unit roots

The ARDL modeling framework relaxes the usual assumption in the cointegration analysis that all variables must be integrated of the same order. However, it is necessary to check the unit root properties of the data series as that method is not valid in the presence of $I(2)$ variables. Here, the integration properties of the data are investigated using two standard unit root tests: the Augmented Dickey-Fuller test (ADF) and the Phillips and Perron (PP) test. For both tests, the null hypothesis is that the data series has a unit root.

From the results presented in Table 2, it is found that all variables used in the study are either $I(0)$ or $I(1)$

which justifies the adoption of an ARDL approach to investigate the presence of cointegration. One can note that the findings are consistent with the earlier studies evoked in the introduction as: (i) the spot price of electricity is found to be mean reverting, (ii) the prices of natural gas are non-stationary, and (iii) the daily consumption of natural gas in each market is also an integrated process of order one. Regarding the spark ratio series, they are also found to be $I(0)$.

The results of the unit root tests conducted with first-differenced series (right panel of Table 2) indicate that none of the variables are integrated of order two. Recall that the absence of a $I(2)$ series is a necessary condition for the application of ARDL-type specifications.

Table 2
Augmented Dickey-Fuller (ADF) and Phillips and Perron (PP) unit root tests.

	Levels				**First differences**			
	ADF		PP		ADF		PP	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
PEG Nord								
q_t	0.21	(0.74)	0.25	(0.75)	-18.94***	(0.00)	-18.99***	(0.00)
s_t	-4.37***	$(0.00)_{C\&T}$	-13.2***	$(0.00)_{C\&T}$	-11.28***	(0.00)	-46.63***	(0.00)
p_t^G	-0.36	(0.55)	-0.37	(0.55)	-23.09***	(0.00)	-23.17***	(0.00)
TRS								
q_t	0,16	(0.73)	0.18	(0.74)	-15.64***	(0.00)	-18.81***	(0.000)
s_t	-4.35***	$(0.00)_{C\&T}$	-12.33***	$(0.00)_{C\&T}$	-10.94***	(0.00)	-44.27***	(0.000)
p_t^G	0.24	$(0.99)_{C\&T}$	0.13	(0.72)	-24.64***	$(0.00)_{C\&T}$	-24.93***	$(0.00)_{C\&T}$
Electricity								
p_t^E	-2.97**	$(0.04)_{C}$	-9.4***	$(0.00)_{C\&T}$	-11.17***	(0.00)	-42.58***	$(0.00)_{C\&T}$

Notes: The table presents the unit root test results for the estimation period: April 1, 2015-December 31, 2016. For the ADF test, the lag structure is the one suggested by the Schwartz information criterion. For the PP test, the truncation lags are decided by Newey-West default. The p-values are provided in parentheses. Here, *C* (respectively *T*) indicates that a constant (respectively a trend) is included in the test equation. Asterisks indicate rejection of the null hypothesis of a unit root at 0.05** and 0.01*** significance levels, respectively.

4.4 Exogeneity

Though the ARDL is known to accommodate endogenous variables, we wish to check the extent to which the next day's consumption level that is released to market participants each day at the beginning of the trading session is a source of endogeneity. In a classical framework, this creates a potential concern, some sort of simultaneous-equation bias as natural gas prices and natural gas consumption would be determined jointly. We use the Wu-Hausman and weak instruments tests of the null hypothesis that the regressors are exogenous. Per the evidence provided in Table 3, we conclude that prices can be considered as exogenous with respect to consumption releases.

Table 3
Endogeneity tests

	df1	df2	Statistic
PEG Nord			
Spark ratio peakload			
Weak instruments	2	402	1.29 ***
Wu-Hausman	1	402	747.10***
Day-ahead gas price			
Weak instruments	2	402	27.84 ***
Wu-Hausman	1	402	619.15 ***
TRS			
Spark ratio peakload			
Weak instruments	2	402	3.40 ***
Wu-Hausman	1	402	751.00 ***
Day-ahead gas price			
Weak instruments	2	402	3.53 ***
Wu-Hausman	1	402	622.00 ***

Notes: Asterisks indicate rejection of the null hypothesis of endogeneity at the 0.01*** significance level.

Again, though endogeneity is not an issue for ARDL estimation, we know that prices are pre-determined and could be considered as exogenous in further applications.

5 Empirical Findings

5.1 Estimation results

We first estimate the linear ARDL specification in equation (1) to determine an appropriate lag structure. In this paper, we adopt the parsimonious one suggested by the Bayesian Information Criterion (BIC) – that is, $p = 1$, $q_1 = 1$ and $q_2 = 1$ – as the associated model residuals show no sign of unmodeled serial correlation. With that lag structure, the short-run dynamics solely considers the instantaneous impacts that the gas price and the spark ratio have on the contemporary demand changes.

We then use that lag structure to estimate our preferred specifications for the NARDL and the TARDL models. For the latter model, the grid search procedure finds that the sum of squared residuals is minimized for $Th = 0.68$ in the PEG Nord market and $Th = 0.62$ in the TRS one. Of particular interest is the relation between these threshold values and the ones obtained from power engineering considerations. After exponentiation, these values correspond to a price of electricity in levels that amount to 1.97 (respectively 1.86) times those of natural gas in the PEG Nord (respectively the TRS) zone. Clearly, these values are commensurate with the heat rates of the large gas-fired power plants installed in France as the nominal thermal efficiencies (i.e., the inverse of the heat rate) of these CCGT plants are in the range $[50\%, 62\%]$.

We now examine whether the presence of short- and long-run asymmetries is supported or not by the data. From the Wald statistics $W_{SR,LR}$ reported in Table 4, we can notice that, whatever the alternative nonlinear model under scrutiny (NARDL or TARDL), the null hypothesis of a fully symmetric ARDL model (i.e., $H_0 : (\theta_1^+ = \theta_1^-)$ and $(\gamma_0^+ = \gamma_0^-)$ and $H_0 : \left(\theta_1^{>Th} = \theta_1^{\leq Th}\right)$ and $\left(\gamma_0^{>Th} = \gamma_0^{\leq Th}\right)$ respectively) is mildly rejected in both the PEG Nord and the TRS zone. Therefore, we further explore whether partial restrictions in the form of either short- or long-run symmetry are supported or not by the data.

The test results in Table 4 convey similar findings for the two nonlinear models and the two markets. We observe low values for the Wald statistics W_{LR} so that we cannot reject symmetry for the coefficients of the spark ratio in the long-run equation. In contrast, the highly significant statistics W_{SR} for the short-run dynamics unambiguously confirm the presence of asymmetry in the short run.

Altogether, these findings reveal: (i) that one should not overlook the presence of asymmetry in the short-run dynamics, and (ii) that it is preferable to opt for symmetric – and thus more parsimonious – specifications for the long-run coefficient associated with the spark-ratio variable. Accordingly, in what follows, our preferred NARDL and TARDL specifications include a symmetric long-run coefficient for that variable and possibly asymmetric ones in the short run.

Our preferred specifications were subjected to several diagnostic tests which included: the presence serial correlation, the presence of heteroscedasticity, and a possible functional misspecification. The results are detailed in Table 5. The autoregressive structures of the estimated models are statistically adequate since there is no evidence of residual autocorrelation (see the Ljung-Box statistic for up to the fifth order). The ARCH test confirms the absence of conditional heteroscedasticity. The Ramsey RESET test for model

Table 4
Asymmetric tests.

	PEG Nord		TRS	
	stat.	p-value	stat.	p-value
NARDL				
$W_{SR,LR}$	3.19	(0.07)*	3.85	(0.09)*
W_{LR}	2.12	(0.14)	0.02	(0.88)
W_{SR}	23.65	(<0.01)***	45.48	(<0.01)***
TARDL				
$W_{SR,LR}$	7.89	(0.05)**	4.60	(0.09)*
W_{LR}	0.15	(0.69)	2.07	(0.15)
W_{SR}	41.67	(<0.01)***	48.25	(<0.01)***

Notes: This table reports the results of a series of Wald tests. The statistic $W_{SR,LR}$ tests the null hypothesis of the restrictions for a fully symmetric ARDL model. The test statistic W_{LR} is for the null hypothesis of long-run symmetry restrictions (respectively $\theta_1^+ = \theta_1^-$ for the NARDL model and $\theta_1^{>Th} = \theta_1^{\leq Th}$ for the TARDL model). The test statistic W_{SR} is for the null hypothesis of short-run symmetry restrictions (respectively $\gamma_i^+ = \gamma_i^-, \forall i$ for the NARDL model and $\gamma_i^{>Th} = \gamma_i^{\leq Th}, \forall i$ for the TARDL model). Asterisks indicate rejection of the null hypothesis at 0.10*, 0.05** and 0.01*** significance levels, respectively.

misspecification based on the powers of the fitted value of consumption shows no sign of functional misspecification for the TARDL model. However, for the NARDL, the p-value of that test is close to the usual 5% significance level in the TRS m arket. To examine the temporal stability of the estimated coefficients, we have also evaluated the cumulative sum of recursive residuals (CUSUM) and CUSUM of square test statistics. The test results are presented in Appendix A (see Fig. 2 and Fig. 3). In all cases, the test statistics are well within the 5% critical bounds which indicates that there is no evidence of parameter instability in any of the models over the estimation period.

The estimation results obtained with our preferred specifications are reported in Table 6 for PEG Nord and in Table 7 for the TRS market. One can note that all models exhibit comparable – and high – explanatory powers.

From the estimates, three relevant series of findings can be d erived. The first series of findings concerns the presence of cointegration. To check the presence of a long-run relationship, we follow the bounds test procedure proposed by Pesaran et al. (2001). In all cases, the F-test statistics reported in the tables systematically exceed the upper bounds critical values at the 1% level which indicates the presence of cointegration among the daily consumption of natural gas, the price of natural gas, and the relative price of electricity and natural gas. Furthermore, the estimated feedback coefficient in the short-run dynamics is

Table 5
Diagnostics test.

	NARDL model		TARDL model	
	Statistic	p-value	Statistic	p-value
PEG Nord				
A: Serial correlation L.-B.(5)	0.006	(0.93)	0.34	(0.55)
B: Heteroscedasticity ARCH(5)	2.66	(0.44)	4.86	(0.18)
C: Functional form RESET(2,3)	1.13	(0.28)	0.79	(0.37)
TRS				
A: Serial correlation	0.22	(0.63)	0.12	(0.72)
B: Heteroscedasticity ARCH(5)	3.63	(0.60)	1.36	(0.92)
C: Functional form RESET(2,3)	3.32	(0.06)*	0.05	(0.81)

Notes: L.-B.(5) is the Ljung-Box Q-statistic for the null hypothesis of no serial correlation up to the fifth order. ARCH(5) is the LM-test for the absence of autoregressive conditional heteroscedasticity with 5 lags. RESET(2,3) is the Ramsey Regression Equation Specification Error Test for model misspecification based on the square and cube of fitted values. Asterisks indicate rejection of the null hypothesis at 0.10*, 0.05** and 0.01*** significance levels, respectively.

negative (as expected) for all models.

Our second series of findings focuses on the estimated long-run multipliers. The positive and statistically significant multiplier obtained for the winter dummy variable is always consistent with our expectations as, unsurprisingly, natural gas demand is larger during the gas heating season. In both markets, the multipliers of the spark ratio variable – which measure the long-term cross-price elasticity of the demand for natural gas to electricity price – are, as expected, positive, and their estimated values are statistically significant at the 1% level. Importantly, we observe that the values of these cross-price elasticities are close to unity and that the values obtained with the TARDL model are slightly lower than the ones obtained with the NARDL one. In contrast, the estimated values for the gas price multiplier – that is, β_2 – are not statistically significant. It is important to keep in mind that this multiplier does not fully capture the long-run influence of natural gas price on gas demand because the long-run price elasticity of natural gas demand is $\beta_2 - \beta_1$. So, to gain further insights on the long-run reaction of gas demand to that price, we report the obtained price elasticities in Table 8 and the associated Wald test. These values are negative, as expected, and the test results confirm the long-run price elastic nature of natural gas demand because, in all cases, the null of a perfectly price inelastic demand is firmly rejected at the 1% level. Moreover, for the TRS market, the values of the demand price elasticity derived from the estimated NARDL and TARDL models are similar and slightly lower than -1.00. In contrast, the values obtained for PEG Nord differ as the long-run demand function associated with the TARDL model is less price elastic than that embedded in the NARDL one.

Table 6
Estimation and test results for the PEG Nord market.

Estimation results		**NARDL**			**TARDL**	
	Estimate	Std. Error	t-stat	Estimate	Std. Error	t-stat
Short-run coefficients						
Constant	1.51	(0.28)	5.40***	1.28	(0.25)	4.90***
Δq_{t-1}	0.04	(0.05)	1.34	0.05	(0.04)	1.66*
Δs^{+}	0.18	(0.03)	5.464***			
Δs^{-}	0.13	(0.03)	3.24***			
$\Delta {s_t}^{>Th}$				0.19	(0.03)	7.25***
$\Delta {s_t}^{\leq Th}$				0.05	(0.04)	1.47
Δp^{G}	0.31	(0.15)	2.13**	0.31	(0.15)	2.18**
$\Delta Winter_y$	-0.001	(0.009)	-0.30	-0.0038	(0.009)	-0.41
$Coint_{t-1}$	-0.07	(0.01)	-5.51***	-0.07	(0.014)	-5.12***
Long-run multipliers						
p_t^G	0.06	(0.23)	0.26	0.34	(0.33)	1.00
s	1.18	(0.11)	10.4***	0.76	(0.099)	7.71***
$Winter_t$	0.74	(0.09)	7.43***	0.83	(0.14)	5.95***
Bounds Test	F-statistic	$I(0)$	$I(1)$	F-statistic	$I(0)$	$I(1)$
W_{coint}	8.12 ***	4.29	5.61	7.25***	4.29	5.61
$Adjust\ R^2$	0.974			0.975		
Observations	405			405		

Notes: The table presents the estimation results for the PEG Nord market for the period April 1, 2015-December 31, 2016. Estimates for the daily dummies are not reported for brevity. Asterisks indicate significance at 0.10*, 0.05** and 0.01*** levels, respectively. The table also reports the non-standard F-test of Pesaran et al. (2001) and the two critical bounds corresponding to the 1% critical level. If that test statistic is lower than the lower bound critical value, the test fails to reject the null of no cointegration. If the test statistic is higher than the upper critical value, the null of no cointegration among the variables is rejected. Asterisks indicate the rejection of the null hypothesis at the 0.01*** level.

Lastly, with regards to short-run dynamics, we observe evidence of asymmetry in the estimated coefficients of the TARDL models. In the two markets, we observe that the instantaneous impact of a change in the spark ratio on gas demand variations is positive and highly significant when the spark ratio is larger than the threshold value. In contrast, the magnitude of that positive impact is less pronounced (and not statistically significant in the PEG market) when the spark ratio is lower than the threshold value. These results are fully consistent with the expectations derived from the technological considerations discussed in section 2. For the NARDL model, the asymmetric coefficients obtained in the two markets are positive and statistically significant. In both markets, a positive change of the spark ratio has a more pronounced instantaneous impact on natural gas demand variations than a negative change.

Overall, our results provide evidence of long-run relationships between natural gas consumption and price

Table 7
Estimation and test results for the TRS market.

Estimation results		**NARDL**			**TARDL**	
	Estimate	Std. Error	t-stat	Estimate	Std. Error	t-stat
Short-run coefficients						
(I)	2.37	(0.30)	7.83***	2.33	(0.32)	7.42***
Δq_{t-1}	0.04	(0.05)	0.90	0.06	(0.05)	1.26
Δs^{+}	0.17	(0.03)	4.99***			
Δs^{-}	0.13	(0.04)	3.64***			
$\Delta {s_t}^{>Th}$				0.18	(0.03)	6.50***
$\Delta {s_t}^{\leq Th}$				0.12	(0.02)	3.61 ***
Δp_t^G	0.38	(0.14)	2.56 *	0.33	(0.15)	2.63**
$\Delta Winter_t$	0.0004	(0.01)	0.06	-0.003	(0.009)	-0.26
$Coint_{t-1}$	-0.11	(0.014)	-7.87***	-0.1	(0.016)	-7.84***
Long-run multipliers						
p^G	0.20	0.20	0.98	0.23	0.27	0.83
s	1.38	0.11	12.8***	1.25	0.12	10.15***
$Winter_t$	0.63	0.09	7.19 ***	0.66	0.11	5.69***
Bounds Test	F-statistic	$I(0)$	$I(1)$	F-statistic	$I(0)$	$I(1)$
W_{coint}	14.03***	4.29	5.61	13.60***	4.29	5.61
$Adjust\ R^2$	0.974			0.975		
Observations	405			405		

Notes: The table presents the estimation results for the TRS market for the period April 1, 2015-December 31, 2016. Estimates for the daily dummies are not reported for brevity. Asterisks indicate significance at 0.10*, 0.05** and 0.01*** levels, respectively. The table also reports the non-standard F-test of Pesaran et al. (2001) and the two critical bounds corresponding to the 1% critical level. If that test statistic is lower than the lower bound critical value, the test fails to reject the null of no cointegration. If the test statistic is higher than the upper critical value, the null of no cointegration among the variables is rejected. Asterisks indicate the rejection of the null hypothesis at the 0.01*** level.

and, with the spark ratio, of elasticities that are of the expected sign and short-run dynamics that are clearly asymmetric, as was expected from the technology considerations above.

5.2 Out of sample analysis

We now use the estimates from our preferred NARDL and TARDL models to generate out-of-sample forecasts. As the NARDL and TARDL models are not nested, it is easy to compare the predictive performance of both models using standard tests. We use the evaluation period covering the year 2017 and the first nine months of 2018.[17]

To gain insight on their predictive accuracy, we use a benchmark formed by the next day's demand forecast

[17] Recall that this period has 234 observations which is large enough to compare the predictions obtained with different models.

Table 8
Long-run price elasticity of natural gas demand.

	PEG Nord			TRS		
	value	stat.	p-value	value	stat.	p-value
NARDL						
$\epsilon^{LR}_{Gas} := (\beta_2 - \beta_1)$	-1.18			-1.10		
$H_0 : (\beta_1 - \beta_2 = 0)$		29.89	<0.01***		108.9	<0.01***
TARDL						
$\epsilon^{LR}_{Gas} := (\beta_2 - \beta_1)$	-0.42			-1.02		
$H_0 : (\beta_1 - \beta_2 = 0)$		67.43	<0.01***		135.17	<0.01***

Notes: The table reports the long-run price elasticities computed from the estimation results. It also reports the result of a Wald test for the null of a price inelastic demand. Asterisks indicate rejection of the null at 0.01*** level.

published each day at 5pm – i.e., directly after the closure of the day-ahead market when closing prices are already known – by infrastructure operators. In the PEG Nord market, GRTGaz is the unique TSO and its forecast can readily be used as a benchmark. In the TRS zone, there are two pipeline operators (GRT Gaz and Teréga) that serve different territories. The forecast of the next day's consumption in the southern zone is thus obtained by summing up the two individual forecasts issued by these two TSOs on their respective websites. To ease comparisons (and obtain error figures measured in energy units), the results presented in this subsection are based on the exponentiated (detransformed) predicted values of the next day's consumption.

Table 9 reports two common measures of accuracy of the predicted values of the next day's consumption: (i) the Root Mean Square Error (RMSE) – measured in GWh – that depends on the scale of the dependent variable, and (ii) the Mean Absolute Percentage Error (MAPE), which is scale independent. Overall, we note that the results obtained for PEG Nord and TRS during the out-of-sample forecast validation period are qualitatively similar. The RMSE statistics obtained for the PEG Nord market are roughly twice the size as the ones obtained for the TRS market. This is consistent with both the relative sizes of these two markets, and the similar magnitudes of the MAPE measures. For both markets, we observe that our TARDL model provides the lowest prediction error statistics and that the accuracy of the forecast based on the NARDL model is comparable to that of the TSOs' forecasts.[18]

As the consumption of natural gas is substantially larger and more variable during the winter season, we

[18] We have also verified that the forecasting performances of the NARDL and TARDL models are better than that of either an AR(1) or a random walk model. The performances of the latter two models are not reported for concision but the results are available from the authors upon request.

also examine the prediction errors obtained for that specific season in our out-of-sample validation period. Unsurprisingly, the RMSE figures obtained during the winter are a bit larger than the ones obtained for the entire evaluation period but the very low MAPE measured during the winter shows that the errors remain modest relative to the average load during that period. Again, the findings indicates that the TARDL model provides the lowest prediction error statistics, which is consistent with the conclusions gained for the entire out-of-sample forecast evaluation period.

In order to statistically compare the predictive accuracy of our models and the forecasts generated by infrastructure operators, we perform the test procedure proposed by Diebold and Mariano (1995) and extended in Harvey, Leybourne, and Newbold (1997). The null hypothesis is that the two forecasting models have the same forecast accuracy and the alternative hypothesis is that model B is more accurate than the reference model A.[19] In Table 9, a negative value for the DM statistic indicates that the second model (model B) delivers significantly better forecasts. In all cases, we observe that model B always delivers forecasts that are significantly more accurate than the reference. In other words, the predictions from the TARDL model significantly outperform both the ones published by TSOs and those from the NARDL model. In addition, one can note that the simpler NARDL model also provides better forecasts than those issued by TSOs. Altogether, these results suggest that the forecasting performance of the tools currently used by TSOs remains poor. Regarding our two specifications, the forecasting accuracy of the TARDL model is the largest, which suggests that one should prefer that model to the NARDL one and thus use the elasticity values obtained with the TARDL model.

The performance of the TARDL model emphasizes the pivotal role of the spark ratio for the purpose of forecasting natural gas demand. This finding confirms the intuition that allowing for a possibly non-zero threshold does help in forecasting the natural gas demand as natural gas consumption does depends on the relative price of electricity to natural gas. Moreover, the good performance of the TARDL relative to the NARDL shows that while the presence of nonlinearity is clearly validated by the data, accounting for this feature alone is not sufficient to outperform the forecasting accuracy of TSOs. We thus conclude that the spark ratio contains much information pertaining to natural gas demand.

[19]The Diebold-Mariano test is known to have poor performance in the case of nested models. It is important to stress that this limitation is not a concern in the present application because the NARDL and TARDL models are not nested.

Table 9
Prediction error statistics values and Diebold-Mariano test statistics.

			NARDL	TARDL	TSOs' forecast
Validation period	PEG Nord	RMSE [GWh]	48.82	**41.56**	47.14
		MAPE [%]	4.3	3.7	**3.1**
	TRS	RMSE [GWh]	26.64	**20.09**	23.76
		MAPE [%]	3.2	**2.9**	5.3
Winter subperiod	PEG Nord	RMSE [GWh]	56.54	**50.04**	65.57
		MAPE [%]	1.8	**1.6**	1.7
	TRS	RMSE [GWh]	36.55	**25.12**	27.59
		MAPE [%]	3.7	**3.0**	3.6
DM Tests	TSO vs. NARDL $\mid$ TSO vs. TARDL	PEG Nord	-15.28 (***)	-17.12 (***)	
		TRS	-5.730 (***)	-17.05 (***)	
	NARDL vs. TARDL	PEG Nord	-16.83 (***)		
		TRS	-16.18 (***)		

Notes: The RMSE and the MAPE are the usual prediction error statistics that are successively evaluated for the entire evaluation period (January 1, 2017 – September 30, 2018) and for the winter subperiod gathering all the observations t in the evaluation period such that $W_t = 1$. The figures in bold indicate that this model has the lowest error value among the three models. DM is the Diebold and Mariano (1995) test of the null hypothesis of no difference in the accuracy of the compared forecasts. Our implementation of that test uses a loss differential defined as the difference between squared forecast errors. Numbers in parentheses are the associated p-values. Asterisks indicate rejection of the null hypothesis at 0.10*, 0.05** and 0.01*** significance levels, respectively.

6 Concluding remarks

Ideally, the prices formed at spot markets for natural gas should reflect the processed information of a large number of market participants. However, anxiety over the liquidity and maturity of some of the European gas hubs has emerged in recent years. The question examined in this paper is therefore whether the information in these day-ahead prices is rich enough to accurately predict the next day's consumption of natural gas?

To answer this question, we examine for the first time the daily interactions between day-ahead prices and daily consumption for two French hubs over the period 2015–2018. Importantly, given the unit-root property of the time series and technological considerations related to the dispatching of gas-fired power plants in the electricity sector, we propose a new nonlinear extension to the ARDL model that permits the response of the natural gas demand to vary with respect to the spark ratio depending on whether that ratio attains or not a certain level: the TARDL model. Our results have shown that its forecasting performance outperforms both those of the usual nonlinear ARDL model – a model that considers the possibly different impacts of positive and negative variations of the spark ratio on gas demand – or those of the tools routinely used by infrastructure operators.

On the whole, our findings have important policy implications, particularly with respect to the quality of the demand forecasts produced by infrastructure operators. The accuracy of these predictions is now emerging as a source of regulatory concern and has recently motivated the adoption of dedicated incentive schemes in some countries (the UK, Italy). Indeed, it has very important implications for: (i) the cost-efficient operation of the gas transportation network, (ii) the quality of the information given to infrastructure users for within-day flow balancing purposes, and (iii) the possibility to use existing gas pipeline infrastructures to supply short-term, linepack-based, flexibility services to a renewable-dominated power sector. Our results suggest that accounting for the information contained in day-ahead prices represents a promising avenue to improve the performance of these demand forecast. That said, the fact that a relatively simple econometric model, solely based on publicly available spot prices, provides a more accurate forecasting procedure certainly points to some deficiencies in the operators' forecasting activities.

Beyond predictive soundness, our research also gives rise to important empirical findings on the economic determinants of the daily demand for natural gas in France. Our results confirm the existence of a long-run relation between the observed demand levels and the spot prices and indicate that this long-run relation is consistent with the conjectures derived from standard microeconomics. Our results indicate that the daily demand has a negative own-price elasticity and a positive cross-price elasticity with electricity. Regarding the short run, our empirical analysis documents the nonlinear nature of the short-run interactions between the observed demand variations and the relative price of electricity to gas. As expected, a positive variation of the spark ratio instantaneously leads to a demand increase but it should be emphasized that the results gained with the TARDL model reveal that its impact is larger when the spark ratio is above a given threshold. Critically, we observe that the value of that threshold is commensurate with the heat rate of the gas-fired power plants.

Notwithstanding the value of our findings, it should be borne in mind that our empirical analysis can be extended in several directions. First of all, it could be interesting to check whether similar results are obtained when applying that methodology to examine the situation prevailing in other European markets. Another strand of research could be the application of that methodology to sectorally disaggregated datasets. Indeed, all consumers do not make optimal demand decisions under the same constraints and do not necessarily demand the same services from natural gas. Hence, the determinants of natural gas demand might differ among different economic sectors. Should disaggregated daily consumption data become publicly available for the various sectors, an analysis of the determinant of the daily demand for natural gas observed in each sector could offer an interesting avenue for future research.

Appendix A:

Table 10
Descriptive statistics for the price of natural gas p_t^G, the electricity price p_t^E and the consumption of natural gas q_t.

	PEG Nord		**TRS**		
	p_t^G	q_t	p_t^G	q_t	p_t^E
Mean	16.61	814.8	17.55	401.8	44.50
Min.	10.65	299.4	11.38	135.5	15.50
Max.	23.16	190.9	25.40	907.5	275.00
Std. dev.	3.32	412.947	3.19	209.52	21.44
Skewness	0.01	0.67	0.09	0.65	4.49
Kurtosis	1.58	2.24	1.76	2.08	39.49
JB	33.28***	31.02***	25.88***	62.31***	24 147***

Notes: The table presents: the mean, min-max, the standard deviation, the skewness, the kurtosis, and the Jarque-Bera test statistics for the series in levels during the model estimation period. Asterisks indicate rejection of the null hypothesis of normality at 0.01*** significance level. For readability, consumption data are measured in GWh.

Figure 2
CUSUM & CUSUMQ test with the NARDL models for PEG Nord (on the left) and TRS (on the right)

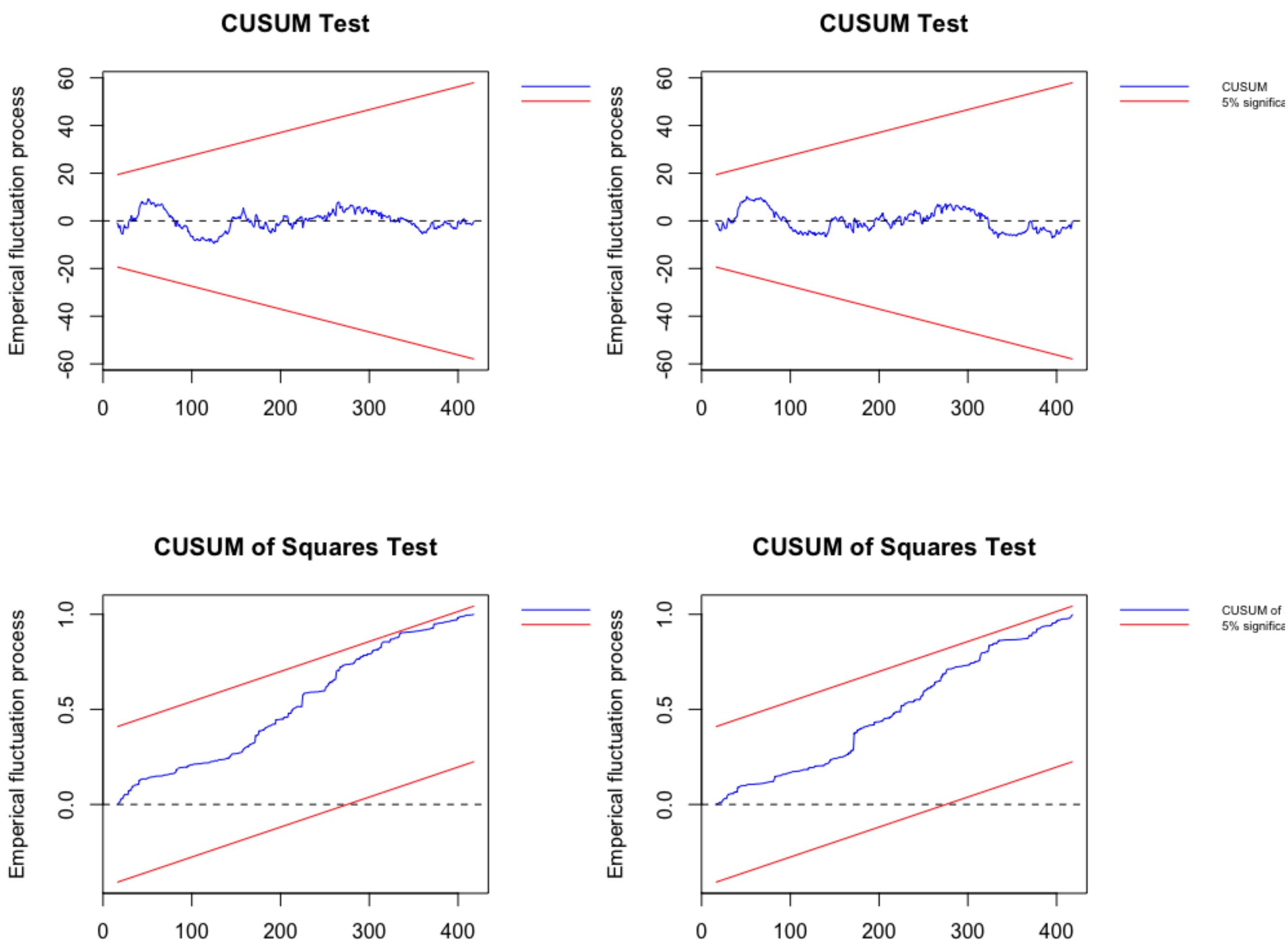
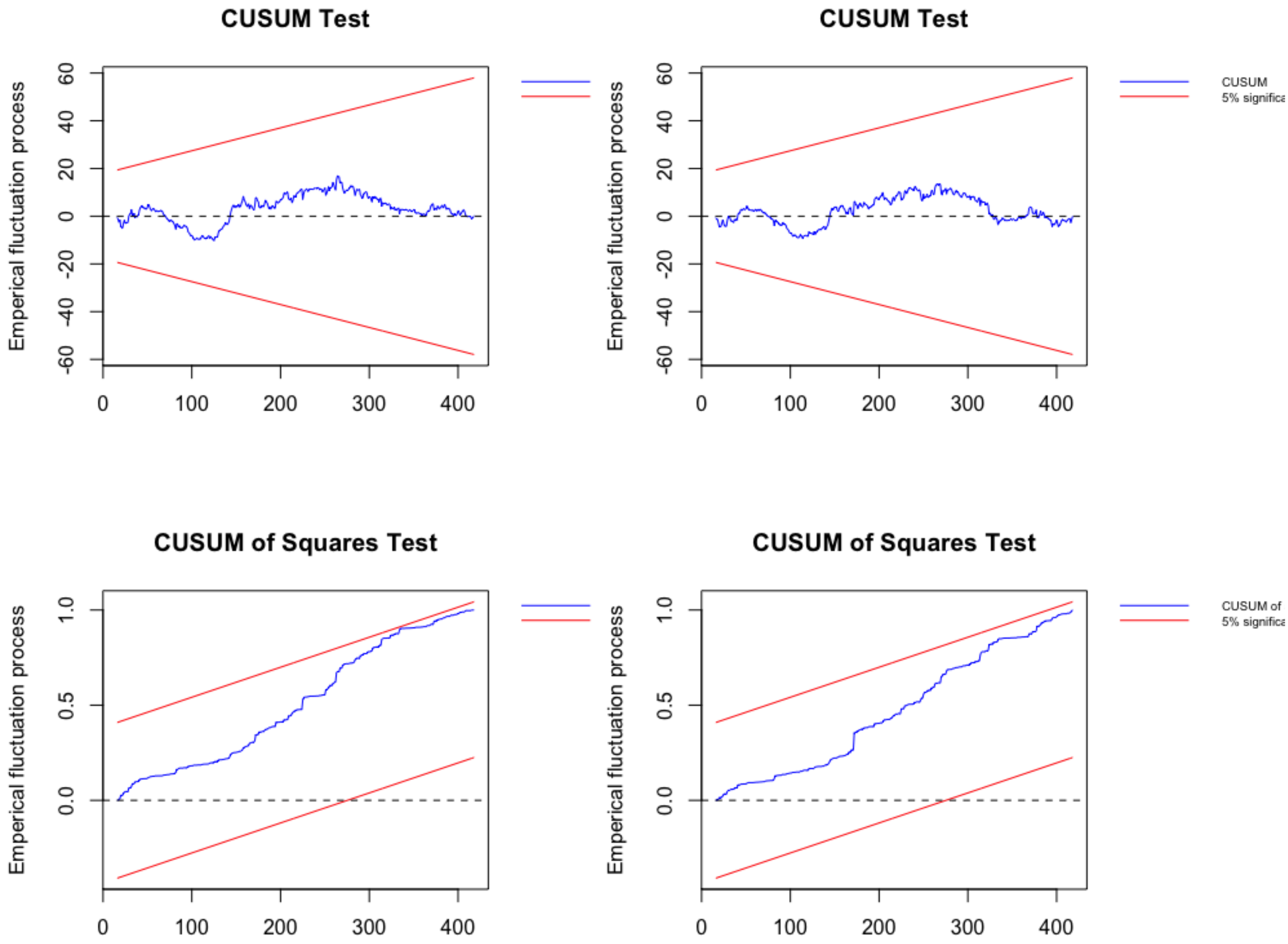


Figure 3
CUSUM & CUSUMQ tests with the TARDL models for PEG Nord (on the left) and TRS (on the right)



References

Abada, I. and O. Massol (2011). "Security of supply and retail competition in the European gas market. Some model-based insights." *Energy Policy* 39(7): 4077–4088.

ACER-ENTSOG (2014). *"ACER-ENTSOG Report on the early implementation of the Balancing Network Code."* European Network of Transmission System Operators for Gas: Brussels.

Arvesen, Ø., Medbø, V., Fleten, S.-E., Tomasgard, A. and S. Westgaard (2013). "Linepack storage valuation under price uncertainty." *Energy*, 52: 155–164.

Bagnai, A. and C.A. Mongeau Ospina (2018). "Asymmetries, outliers and structural stability in the US gasoline market." *Energy Economics*, 69: 250–260.

Bordignon, S., Bunn, D.W., Lisi, F. and F. Nan (2013). "Combining day-ahead forecasts for British electricity prices." *Energy Economics*, 35: 88–103.

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

CRE (2012). "*Délibération de la Commission de Régulation de l'Énergie du 13 décembre 2012 portant décision sur le tarif d'utilisation des réseaux de transport de gaz naturel*." Commission de Régulation de l'Énergie: Paris.

De Menezes, L. and M.A. Houllier (2016). "Reassessing the integration of European electricity markets: A fractional cointegration analysis." *Energy Economics*, 53: 132–150.

De Vany, A. and W.D. Walls (1993). "Pipeline Access and Market Integration in the Natural Gas Industry: Evidence from Cointegration Tests." *The Energy Journal*, 14(4): 1–19.

Diebold, F.X. and R. Mariano (1995). "Comparing predictive accuracy." *Journal of Business Economics and Statistics*, 13: 253–265.

Engle, R. F. and Granger, C. W. J., (1987). "Co-Integration and Error Correction: Representation, Estimation, and Testing". *Econometrica*, Vol. 55, No. 2. (Mar., 1987), pp. 251-276.

ENTSOG (2017). "*Balancing Network Code. Implementation and Effect Monitoring Report*." European Network of Transmission System Operators for Gas: Brussels.

Forbes, K. and Zampelli, E., (2014) "Do Day-Ahead Electricity Prices Reflect Economic Fundamentals? Evidence from the California ISO." *The Energy Journal*, 35(3): 129–144.

Geweke, J. (1978) "Temporal Aggregation in the Multiple Regression Model." *Econometrica*, 46(3): 643–661.

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

Giulietti, M., Grossi, L. and M. Waterson (2012). "A Rough Analysis: Valuing Gas Storage" *The Energy Journal*, 33(4): 119–141.

Gopalakrishnan, A. and L.T. Biegler (2013). "Economic Nonlinear Model Predictive Control for periodic optimal operation of gas pipeline networks." *Computers & Chemical Engineering*, 52: 90-99.

Green, R. and N. Vasilakos (2010). "Market behaviour with large amounts of intermittent generation." *Energy Policy*, 38(7): 3211-3220.

Greenwood-Nimmo, M., Shin, Y., and T. van Treek (2011). "*The Asymmetric ARDL model with multiple unknown thresholds decompositions an application to the Phillips curve in Canada*." mimeo: http://www.greenwoodeconomics.com/(MG2)IMK_multi.pdf.

Hallack, M. and M. Vazquez (2013). "European Union regulation of gas transmission services: Challenges in the allocation of network resources through entry/exit schemes." *Utilities Policy*, 25: 23–32.

Harvey, D., Leybourne, S. and P. Newbold (1997). "Testing the equality of prediction mean squared errors.",*International Journal of Forecasting*, 13 (2), 281-291.

Heather P. and B. Petrovich (2017). "*European traded gas hubs: an updated analysis on liquidity, maturity and barriers to market integration*." Oxford Institute for Energy Studies. Energy Insight: 13 May 2017.

Hoel, M. and S. Strøm (1987). "*Supply security and import diversification of natural gas.*" In R. Golombek, J. Vislie, M. Hoel (Eds.), "Natural Gas Markets and Contracts", North-Holland.

Johansen, S. and K. Juselius (1990). "Maximum Likelihood Estimation and Inference on Cointegration with Applications to Demand for Money" *Oxford Bulletin of Economics and Statistics* 52, 169-210.

Karakatsani, N. and D.W. Bunn (2008). "Forecasting electricity prices: the impact of fundamentals and time-varying coefficients." *International Journal of Forecasting*, 24: 764–785.

Keyaerts, N., Hallack, M., Glachant, J.-M. and W. D'haeseleer (2011) "Gas market distorting effects of imbalanced gas balancing rules: Inefficient regulation of pipeline flexibility." *Energy Policy*, 39(2): 865–876.

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

Lucia, J. and E. Schwartz (2002). "Electricity prices and power derivatives: evidence from the Nordic power exchange." *Review of Derivative Research*, 5: 5–50.

Manne, A.S., Roland, K. and G. Stephan (1986). "Security of supply in the Western European market for natural gas." *Energy Policy*, 14(1): 52–64.

Markandya, A. and M. Pemberton (2010). "Energy security, energy modeling and uncertainty." *Energy Policy*, 38(4): 1609–1613.

Miriello, C. and M. Polo (2015). "The development of gas hubs in Europe." *Energy Policy*, 84: 177–190.

National Grid (2018) "*National Grid Gas (NTS) System Operator Incentives. 2018/19 Supporting Information*." October 2018. UK.

Pal, D. and S.K. Mitra (2015). "Asymmetric impact of crude price on oil product pricing in the United States: An application of multiple threshold nonlinear autoregressive distributed lag model." *Economic Modelling*, 51: 436–443.

Pesaran, M.H. and Y. Shin (1998). "*An Autoregressive Distributed-led Modelling Approach to Cointegration Analysis*." Cambridge: Cambridge University Press.

Pesaran, M.H., Shin, Y. and R. Smith (2001). "Bound testing approaches to the analysis of level relationship." *Journal of Applied Econometrics*, 16: 289–326.

Qadrdan, M., Chaudry, M., Wu, J., Jenkins, N. and J. Ekanayake (2010). "Impact of a large penetration of wind generation on the GB gas network." *Energy Policy*, 38(10): 5684–5695.

Ramanathan, R., Engle, R., Granger, C.W.J., Vahid-Araghi, R. and C. Brace (1997). "Short-run forecasting of electricity loads and peaks." *International Journal of Forecasting*, 13: 161–174.

Renou-Maissant, P. (2012). "Toward the integration of European natural gas markets: A time-varying approach." *Energy Policy*, 51: 779–790.

Serletis, A. and J. Herbert (1999). "The message in North American energy prices." *Energy Economics*, 21(5): 471–483.

Shin, Y., Yu, B. and M.J. Greenwood-Nimmo (2014). "*Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework*". In William, C. Horrace, Robin, C. Sickles (Eds.), "'Festschrift in Honor of Peter Schmidt"', Springer Science & Business Media, New York.

Sun, X., Huang, D. and G. Wu (2012). "The current state of offshore wind energy technology development." *Energy*, 41: 298–312.

Thoenes, S. (2014). "Understanding the Determinants of Electricity Prices and the Impact of the German Nuclear Moratorium in 2011."*The Energy Journal*, 35(4): 61–78.

Tran, T.H., French, S., Ashman, R. and E. Kent (2018). "Linepack planning models for gas transmission network under uncertainty." *European Journal of Operational Research*, 268(2): 688–702.

Uría, R. and J. Williams (2007). "The Supply of Storage for Natural Gas in California." *The Energy Journal*, 28(3): 31–50.

Worthington, A., Kay-Spratley, A. and H. Higgs (2005). "Transmission of prices and price volatility in Australian electricity spot markets: a multivariate GARCH analysis." *Energy Economics*, 27(2): 337–350.