

A joint model for temperature and natural gas with an application to the US market

Roberto Baviera*

Teodoro Federico Mainetti†

October 5, 2016

Abstract

Natural gas spot prices and temperatures have been studied in detail in the literature as separate processes. We propose a simple joint model that, in spite of its parsimony, describes accurately many *stylized facts* of the two time series: in particular we show the role played by a time-delay parameter in order to take into account the impact of temperature forecast in cross-dependency. We discuss in detail a stepwise procedure in order to calibrate model parameters, describing the elementary estimation techniques involved and the statistical accuracy achieved. In the analysis we focus on the benchmark market in the U.S.A. (Henry Hub) and the temperatures in the Northeast and Midwest regions; we observe a negative, statistically significant, gas-temperature correlation in the cold season.

Keywords: Seasonality, NIG, temperature forecast, joint model, statistical bootstrap.

JEL Classification: C15, C22, C51.

Address for correspondence:

Roberto Baviera
Department of Mathematics
Politecnico di Milano
32 p.zza Leonardo da Vinci
I-20133 Milano, Italy
Tel. +39-02-2399 4630
Fax. +39-02-2399 4621
roberto.baviera@polimi.it

*Department of Mathematics, Politecnico di Milano, 32 p.zza Leonardo da Vinci, I-20133 Milano, email: roberto.baviera@polimi.it

†ATIE Uno, 30 via Macon, I-23900 Lecco, email: teodoro.mainetti@unoi.it

A joint model for temperature and natural gas with an application to the US market

1 Introduction

It sounds quite natural to consider gas spot prices weather sensitive because demand for gas depends on temperatures and supply is constrained. In this study our main goal is to introduce a joint model for natural gas and temperatures in the U.S.A.

According to the Residential Energy Consumption Survey of the U.S. Energy Information Administration (EIA), in 2009, the latest year with available data, 48% of energy consumption in U.S. homes was for heating and cooling, down from 58% in 1993 (EIA 2013a). A major effect can be observed in the cold season: the U.S. consumes almost twice as much natural gas (also NG hereinafter) in winter than in summer. Furthermore, this increase is quite localized geographically: during the years 2001 – 2014 on average 2/3 of the total increase is due to Northeast and Midwest, and in some years these two regions account for more than 80% of this increase (EIA 2015). Thus, colder temperatures in these U.S. regions can lead to demand shocks due to heating.

Moreover, even if in a less pronounced way, temperatures can also affect gas prices during the cooling season since 87% of U.S. households are equipped with air conditioning according to the same survey. About 30% of U.S. electricity was generated by natural gas: warmer temperatures can increase the demand for air conditioning which in turn increases the power sector's demand for natural gas (EIA 2014).

Temperatures and natural gas prices have been studied in detail as separate underlyings. The dynamics of daily average temperatures has been analyzed in recent years; the two most known approaches are the one proposed by Benth and Benth (2007, 2012) and by Campbell and Diebold (2005). Both studies share several similarities. Temperatures have a trend, a seasonal and an autoregressive (AR) component; residuals are Gaussian random variables. We follow the first approach for two main reasons. On one hand for parsimony: the model in Benth and Benth (2007) is sophisticated enough to explain main *stylized facts* in temperatures' dynamics just as well as the parameter intensive alternative. On the other hand for calibration purposes: the approach of Benth and Benth (2007) allows a simple stepwise procedure which is more in line with standard practitioners' behaviour and useful when introducing the joint model.

Natural gas spot price dynamics can be modelled following several approaches (see e.g. Benth *et al.* 2008). One of the simplest models able to capture a seasonally varying level, mean reversion and the frequent occurrence of spikes, has been introduced in Benth and Benth (2004). This model is a generalization of the continuous-time Schwartz model (Schwartz 1997) with non Gaussian residuals. The model for temperatures we consider in this study has been described in the discrete-time case by Benth *et al.* (2008): the deseasonalized logarithmic spot price is an AR(1) with Normal Inverse Gaussian (NIG) residuals that are able to model the heavy-tailed spot price innovations (Barndorff-Nielsen 1997). Our contribution to the existing literature is the proposal of a joint process for temperature and natural gas that can be easily calibrated on temperatures and market data.

This study is also related to the literature on temperatures' impact on prices. The question is whether temperature forecast could be a good predictor for commodity prices. This question

has been addressed in a different commodity market: the agricultural one. Roll study (Roll 1984) is the most well known example: he wondered whether frozen-concentrated orange juice (FCOJ) prices were related to weather forecasts at a single location in the U.S.A. At the time of the study, FCOJ had also a geographically concentrated production, with more than 98% of U.S. oranges' production in the central Florida region around Orlando. Cold temperature can have a significant influence on FCOJ production and can induce a supply shock. However Roll concludes that “temperature has no remaining predictive content” (see Roll 1984, p. 879).

This finding has been a famous confirmation of the efficient markets hypothesis, i.e. prices reflect all available information: the lack of relation between prices and temperatures in such a simple market as the one of orange juice has led many researchers to consider unlikely that a relation exists in more liquid commodity markets; studies in electric power markets are probably the main exception due to the limited storability of electricity. Some authors have recently proposed temperature as one of the explicatory variables in order to forecast prices (see e.g. Weron 2014, for a review).

There are very few studies in the literature that model the dependency of natural gas spot prices and temperatures in a discrete time dynamics. Recently Stoll and Wiebauer (2010) have proposed a model that relates the natural gas price at a European hub, the Title Transfer Facility in the Netherlands, to the cumulated Heating Degree Days (HDD) measured in a German meteorological station, the one located in Hannover. Their model incorporates in gas price a temperature component via HDD, an average cumulated temperature (below a threshold) over a winter (beginning on the 1st of October), as a proxy of the filling level of gas storage; their approach has been generalized by Müller *et al.* (2015), who include also oil price as another exogenous factor in a model for Central Europe gas prices.

Our approach is quite different from Stoll and Wiebauer (2010) and from Müller *et al.* (2015). We introduce a complete joint model for the two daily dynamics of Henry Hub natural gas prices and temperatures in some U.S. cities in the Northeast and Midwest regions; we also discuss in detail the significance of model parameters calibrated on real data with a particular focus on temperature forecast.

The remaining part of the paper is organized as follows. In section 2 we introduce the dataset we have considered. In section 3 we briefly recall the models for temperatures and natural gas and introduce the correlation and the delay parameters that characterize the dependency, while in section 4 we discuss the calibration technique, the statistical significance of the estimated quantities and the robustness of the analysis. Finally in section 5 we state some concluding remarks.

2 Dataset

The dataset we have analyzed includes both natural gas data and temperatures.

Natural gas prices are location based prices. Henry Hub (commonly referred to as HH) spot price is natural gas benchmark in the U.S.A. (also called backbone) and prices at other locations are derived by adding “bases” to the Henry Hub price. HH prices are based on next day delivery at the Henry Hub in Louisiana and they are in US Dollars per Million Btu (\$/MMBtu). Official daily closing prices are established at 2:30 p.m. in the trading floor of the New York Mercantile Exchange (NYMEX). The spot market is traded every business day and settled on the next business day and non-business days if there are any before the next business day. For instance, natural gas traded on July 19, 2012 (Thursday) is settled on July 20, 2012 (Friday), and natural gas traded on July 20, 2012 is settled on July 21 (Saturday), July 22 (Sunday), and July 23, 2012 (Monday).

Financial data are provided by Bloomberg: we consider HH natural gas spot daily prices (NGUSHHUB Index) for more than 20 years since March 10, 1994 up to December 31, 2014. In figure 1 we plot the time series of HH natural gas spot prices. The time series includes the HH closing price at each business day in the lag a part from August 29, 2005 (due to Hurricane Katrina and the flooding of New Orleans) and the period between September 23, 2005 and October 6, 2005 included (due to Hurricane Rita). Also prices for September 2, 2008 (due to Hurricane Gustav) and September 15, 2008 (due to Hurricane Ike and Lehman default) are missing from the dataset, resulting in 5203 price recordings.

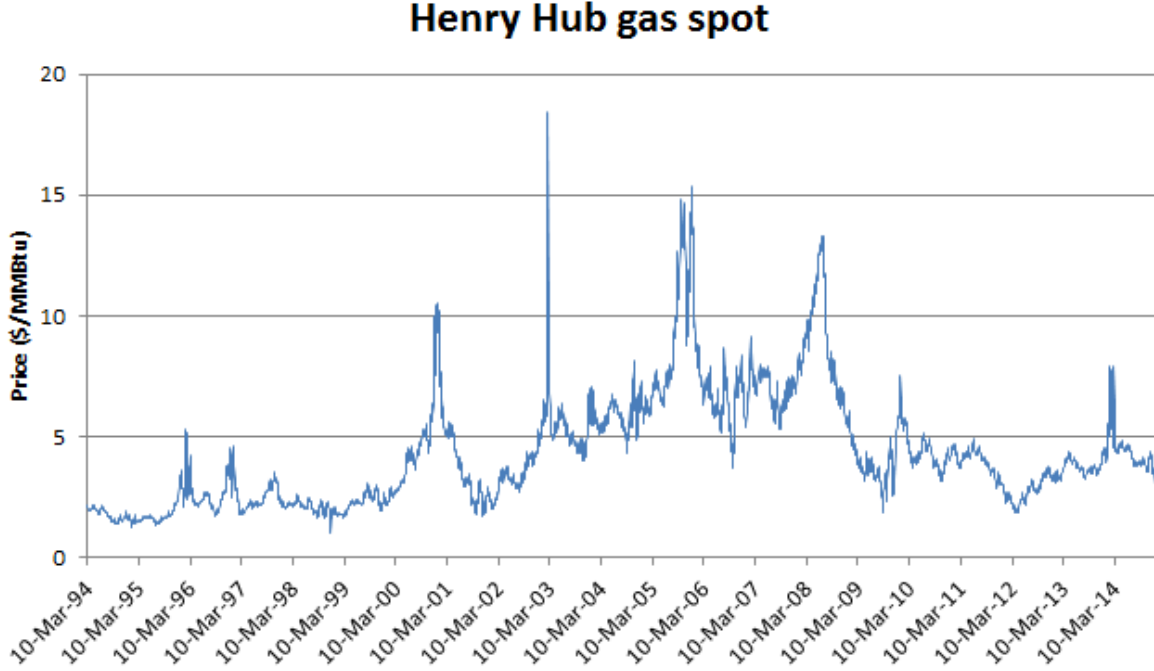


Figure 1: *Henry Hub natural gas spot prices in the time window analyzed: March 10, 1994 up to December 31, 2014*

The time series we have analyzed is complete. Natural gas began trading on the NYMEX on April 3, 1990; however a truly liberalization in the market began only after Federal Energy Regulatory Commission Order 636 (April 8, 1992). This order fostered competition, since it required pipelines to unbundle their services, ending the pipelines' traditional middleman role and converting them to transportation companies. For this reason, only since 1994 natural gas has become a liquid market; before 1994 closing prices remained unchanged for several days and in some cases even for weeks.

In our analysis we consider also natural gas deliveries to all consumers (residential, commercial, industrial sectors, vehicle fuel and electricity generation) during the period between January 2001 and November 2014. These monthly data refer to the amount of gas delivered to end users divided for each state in the U.S.A.: this dataset is provided by EIA; unfortunately, consumption data have not been registered by EIA before 2001.

Temperature data are provided by the National Climatic Data Center. We consider the daily mean temperatures of three cities in the U.S.A. in the Northeast region (New York, Boston) and in the Midwest region (Chicago). As we have discussed in the introduction, Northeast and Midwest regions play a crucial role in the natural gas market: moreover, New York is the largest demand center in the U.S.A., Boston is the location in the Northeast farthest from gas

sources (Gunnarshaug and Ellerman 1998) and Chicago the most important demand center in the Midwest region. Even if data in many locations are registered with the same accuracy since the beginning of 1949, we just focus on the same lag of natural gas data, so, except where differently stated, this constitutes our sample period.

A comment can be useful on the filtering techniques we have adopted on empirical data. As in Benth *et al.* (2008) we have filtered out the 29th of February in each leap year; we have also eliminated 6 values in HH time series on the day after Thanksgiving: natural gas spot market was closed on those dates and the registered price was just equal to previous day value. Moreover, in order to avoid any smoothening in the dataset, we have preferred not to replace missing observations for natural gas prices with the average of the two closest nearby recordings as in Benth *et al.* (2008); in any case, as discussed above, in the natural gas data series there are just 13 missing values and then the impact is negligible whatever is the chosen rule for treating these data.

3 The model

In this section we present the joint model for temperature and natural gas. Three are the main objectives of this study: i) obtain a parsimonious joint model, ii) propose a stepwise procedure for model calibration, iii) provide statistical significance of the calibrated parameters.

The reason for parsimony is rather intuitive and due not only to simplicity; having a not so large set of parameters looks quite natural in order to avoid over-fitting and to obtain statistically significant quantities.

We also require that the model could be calibrated via a stepwise procedure. As explained in Benth *et al.* (2008) and Benth and Benth (2012), it is crucial to estimate separately model components as in the standard methodology followed by practitioners: examining step-by-step intermediate components “one gets a good insight in data structure and is likely able to come up with a precise model” (see Benth and Benth 2012, p. 593). In Figure 2 we show the main steps in the proposed joint model. First we analyze separately natural gas and temperature models in their seasonal, AR and residual components and then we consider the complete joint model with the two key parameters of correlation and delay. We first briefly recall the natural gas model (see subsection 3.1) and in subsection 3.2 the temperature model presented in Benth and Benth (2007); finally in subsection 3.3 correlation and delay are introduced as main parameters for the joint model. **Proposition 1** in subsection 3.3 relates both parameters to natural gas and temperature residuals.

3.1 A model for natural gas

The model for natural gas is very similar to the one described in Benth *et al.* (2008), where deseasonalized spot prices follow an AR process with NIG residuals. This discrete model corresponds to the continuous one introduced in Benth and Benth (2004) for the National Balancing Point UK market

$$\begin{cases} S(t_b) = \Lambda^S(t) \exp(x(t_b)) \\ x(t_b) = b_1 x(t_b - 1) + \Delta I(t) \\ \Delta I(t) \sim NIG(\alpha, \beta, \delta) \text{ i.i.d.} \end{cases} \quad t_b \in \mathbb{N} \quad (1)$$

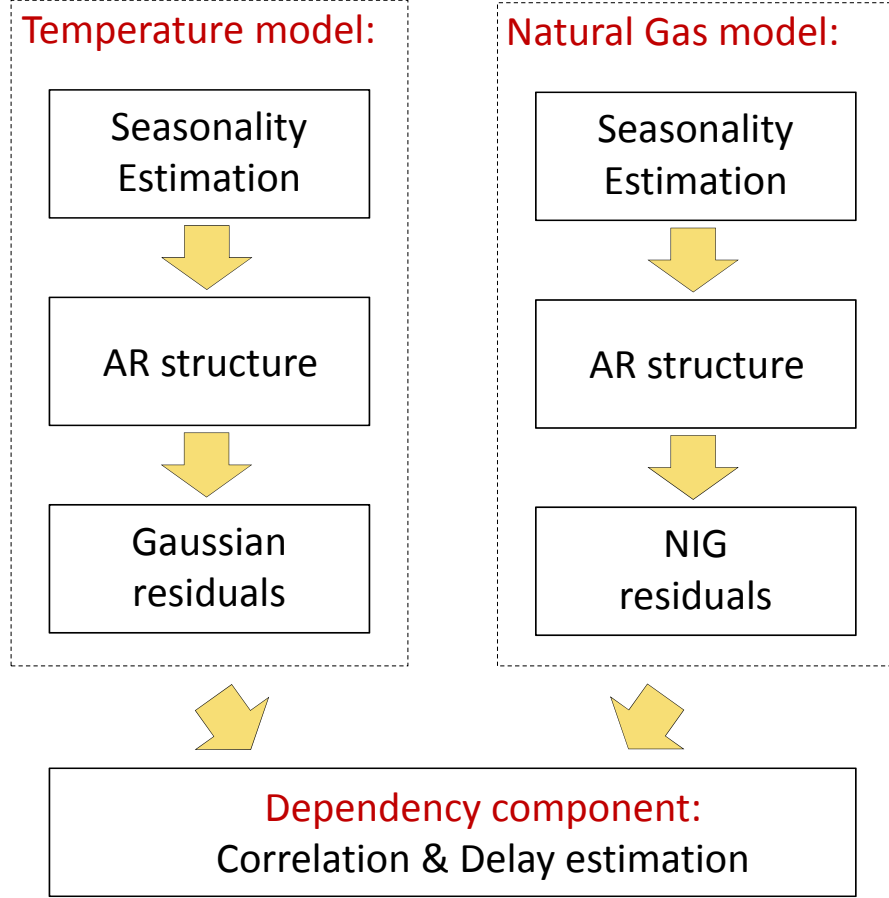


Figure 2: *Main steps in the proposed joint model. First temperatures' and natural gas' models are estimated separately, then main model parameters for dependency (correlation and delay) are evaluated from residuals.*

where t_b is sequence's business time, while t is a short notation for $t(t_b)$ and indicates the calendar time associated to the business time t_b . The seasonal mean function for natural gas is

$$\Lambda^S(t) = a_0^S + \sum_{k=1}^p [a_{2k-1}^S \cos(2k\pi t/365) + a_{2k}^S \sin(2k\pi t/365)] \quad \text{with} \quad p = 2. \quad (2)$$

Two are the main differences of the model for natural gas we consider w.r.t. the one described in Benth *et al.* (2008): i) the form of the seasonality function and ii) the calibration of residuals with three parameters' NIG.

The seasonal mean function adopted in this study (2) differs from Benth *et al.* (2008) in three major aspects. First, we have not included any drift term since no trend has been observed in our natural gas series; second, we have considered the seasonality on the calendar time and not on the business time, since the seasonal behaviour of prices is related to the true calendar days; third, we have considered both an annual and a semiannual seasonality: as discussed in section 1 we expect to find two local maxima related to the heating and cooling seasons.

The seasonal mean function with $p = 2$ describes accurately the observed seasonality on gas prices. In figure 3 we plot the average price of each calendar day for the dataset analyzed and

we observe that its seasonality with two periods with higher prices, one more pronounced in the winter season and another in the summer season, is well described by equation (2). We have also considered the case with $p = 1$ (dashed line) as in Benth *et al.* (2008): this description does not look adequate in the U.S. case, with the maximum on the 23rd of January, quite different from the observed maximum in prices.

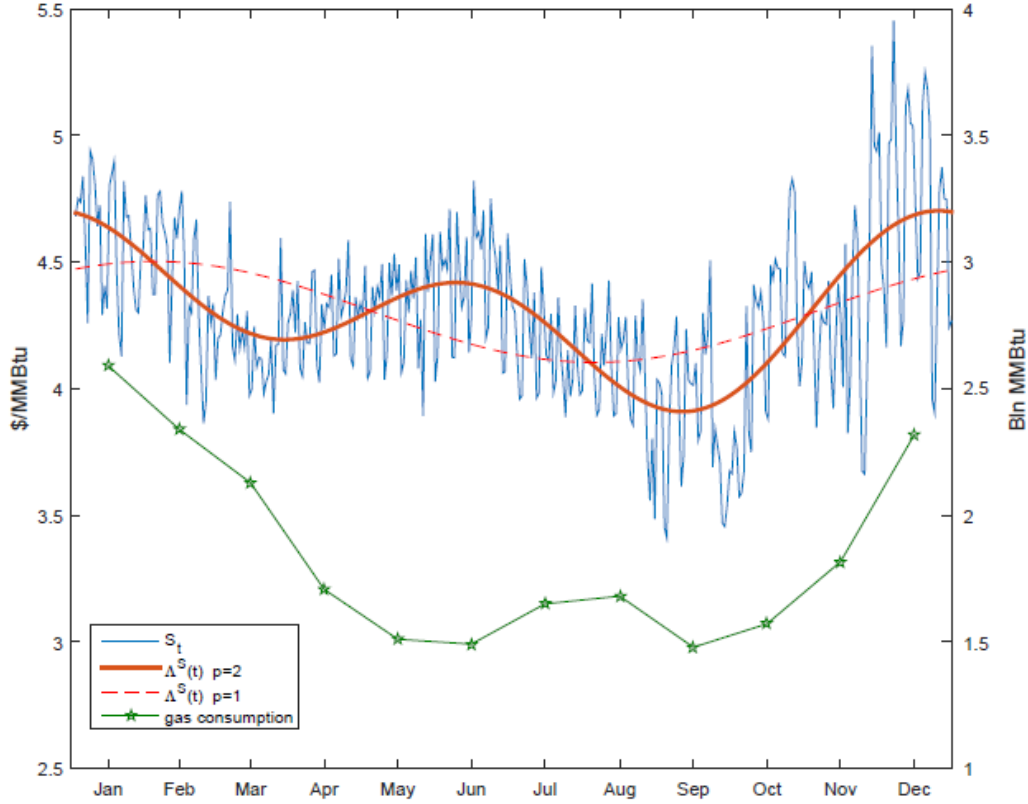


Figure 3: Average daily Henry Hub gas price (blue line) from March 10, 1994 to December 31, 2014 and seasonality $\Lambda^S(t)$ (red) on the left scale in \$/MMBtu. We observe the two local maxima in the seasonality, one on the 26th of December in the winter season and the other one, lower than the previous one, on the 9th of June during the summer period; with a dashed line we have also plotted the seasonality with $p = 1$ as in Benth *et al.* (2008). The green line is the average monthly natural gas consumptions in the U.S.A. in the period 2001-2014 on the right scale in Billion MMBtu.

Even if, as we shall discuss in the next section, three out of five parameters are not significant at the 10% level, the seasonal mean function of temperature seems to capture reasonably well the observed seasonality with two maxima on the 26th of December and 9th of June.

For this reason, in order to show that the calibrated seasonality is in line with market prices behaviour we have also analyzed monthly consumption data from EIA. In figure 3 we show that the average monthly consumption in the period 2001 – 2014 follows a seasonality pattern similar to prices, with two maxima on January and July-August. It looks reasonable to observe maxima in prices approximately the month before maxima reported for consumption.

Since outliers may influence seasonality analysis, they have been removed following the same technique described in Benth *et al.* (2008). Following that technique 84 outliers have been removed and the seasonality has been obtained via a least squared minimization. After estimating the seasonality, outliers have been inserted back into the time series.

In the de-seasonalized time series we have estimated the autoregressive coefficient b_1 with the least squared minimization and NIG model parameters via a maximum likelihood estimation. Another difference w.r.t. Benth *et al.* (2008) is that we do not consider all four parameters $NIG(\mu, \alpha, \beta, \delta)$ but only three because the parameter μ is not significant at 10% level (see also the discussion in the next section), we impose $\mu = 0$ in order to simplify the model.

3.2 A model for temperatures

The model for temperatures is similar to the one introduced by Benth and Benth (2007)

$$\begin{cases} T(t) = \Lambda^T(t) + y(t) \\ y(t) = c_1 y(t-1) + c_2 y(t-2) + c_3 y(t-3) + \Delta J(t) \\ \Delta J(t) \sim N(0, \sigma_T^2(t)) \text{ i.i.d.} \end{cases} \quad t \in \mathbb{N} \quad (3)$$

where the seasonal mean function for temperatures is

$$\Lambda^T(t) = a_0^T + a_d^T t + [a_1^T \cos(2k\pi t/365) + a_2^T \sin(2k\pi t/365)] . \quad (4)$$

This model looks quite similar to the one for natural gas. The seasonal mean function of temperature $\Lambda^T(t)$ incorporates, besides a seasonal component describing periodic temperature variations over the year, the level of the long-term average temperature a_0^T and a trend (or drift component) $a_d^T t$ related to the observed risen in temperatures due to global warming and urbanisation. As already observed in Benth and Benth (2007) the seasonal behaviour is well captured by a yearly periodicity.

In figure 4 we show last 5 years temperatures measured at the New York La Guardia weather station and the seasonality function obtained via least squares minimization. As in the natural gas case the AR structure is modelled after removing the seasonal mean function.

Analogously to the European case considered by Benth, the volatility presents a seasonal form. It can be calibrated, either in the same way as in Benth and Benth (2007) via

$$\sigma_T^2(t) = a_0^\sigma + \sum_{k=1}^p [a_{2k-1}^\sigma \cos(2k\pi t/365) + a_{2k}^\sigma \sin(2k\pi t/365)] \quad (5)$$

or considering simply the average monthly values. In equation (5) p can be equal either to 3 (New York and Boston) or to 4 (Chicago): next terms in Fourier expansion are not significant in the time series analyzed.

In Figure 5 we show the comparison between these two approaches. Both are a good approximation of squared residuals. However, in order to obtain a reliable calibration using equation (5) we need a large amount of data. Due to the limited time series we consider, we choose the average monthly values for volatility.

3.3 Dependency between natural gas prices and temperatures

In previous subsections we have described the model for both temperature and natural gas as separate time series via a seasonal component, an AR component and a residual term. These residuals at each time $t \in \mathbb{N}^+$ can be written as

$$\begin{cases} \Delta J(t) = \sigma_T(t) \epsilon^T(t) \\ \Delta I(t) = \beta z(t) + \sqrt{z(t)} \epsilon^S(t) , \end{cases} \quad (6)$$

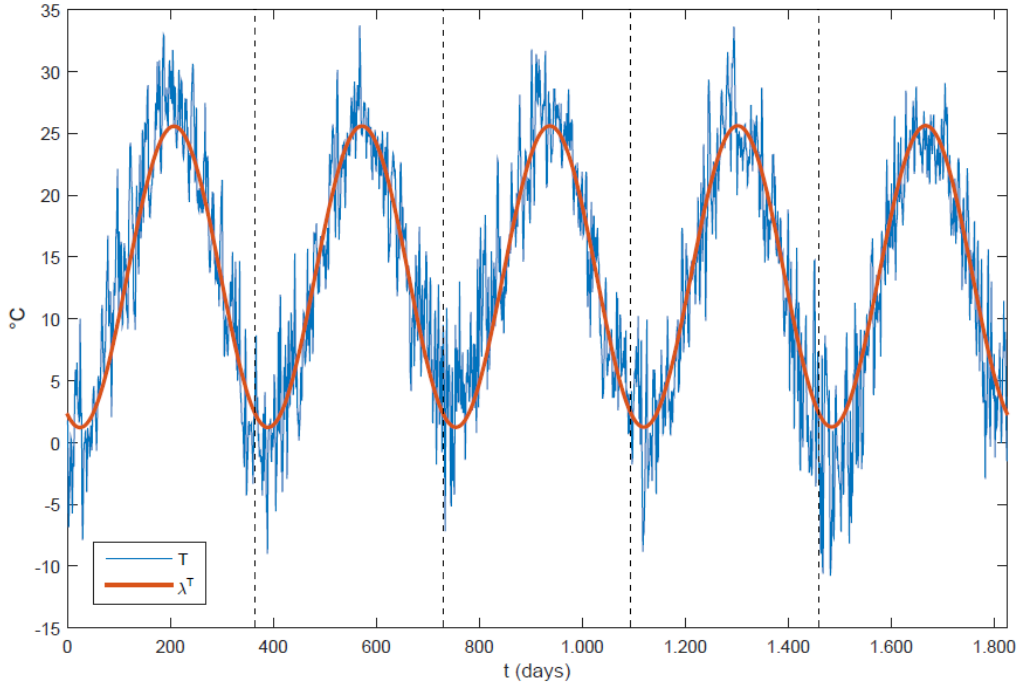


Figure 4: *New York La Guardia temperatures in degrees centigrades over the period January 1, 2010 - Dececeber 31, 2014. The red line describes the seasonal mean function of temperature $\Lambda^T(t)$: a yearly periodicity seems to describe well the main seasonal component in temperature series.*

where $\epsilon^S(t)$ and $\epsilon^T(t)$ are two standard normal variables, while $z(t)$ is an IG mixing variable independent from the other two.

The simplest joint model can be introduced via a correlation function $\rho(t)$ and a delay parameter τ

$$\mathbb{E} [\epsilon^T(t') \epsilon^S(t)] = \rho(t) \delta(t', t + \tau) \quad \tau \in \mathbb{N}^+ \quad (7)$$

where $\delta(n, m)$ is a standard Kronecker delta between two numbers $n, m \in \mathbb{N}$.

This dependence is rather unusual in financial markets. In autoregressive models financial quantities depend on past values; in this case we state that gas prices depend on future values for temperatures, since τ is a non negative natural number.

The fact that not yet observed temperatures affect natural gas prices could appear quite strange at a first glance. Few words should be spent to comment this passage which is crucial in our model construction.

In the nineties weather forecasts have improved significantly thanks to the enhanced processing power of computational resources available. For example, one of the primary sources of weather forecast in the U.S.A., the National Weather Service (NWS), in 1994 started providing weather updates every 3 hours (and hourly updates after 1998). Prior to 1994 these updates were just twice a day. Furthermore, many commodity traders also subscribe to private forecasting services that could provide more precise forecasts than the NWS in some specific locations.

We do not know which has been the most accurate forecasting service in the time window analysed in this study (but we can guess that a commodity trader knows this piece of information, if it is relevant in price formation) and we have not his/her temperatures forecasts: for the analysis we are interested in, the measured temperature is the best estimator of future days

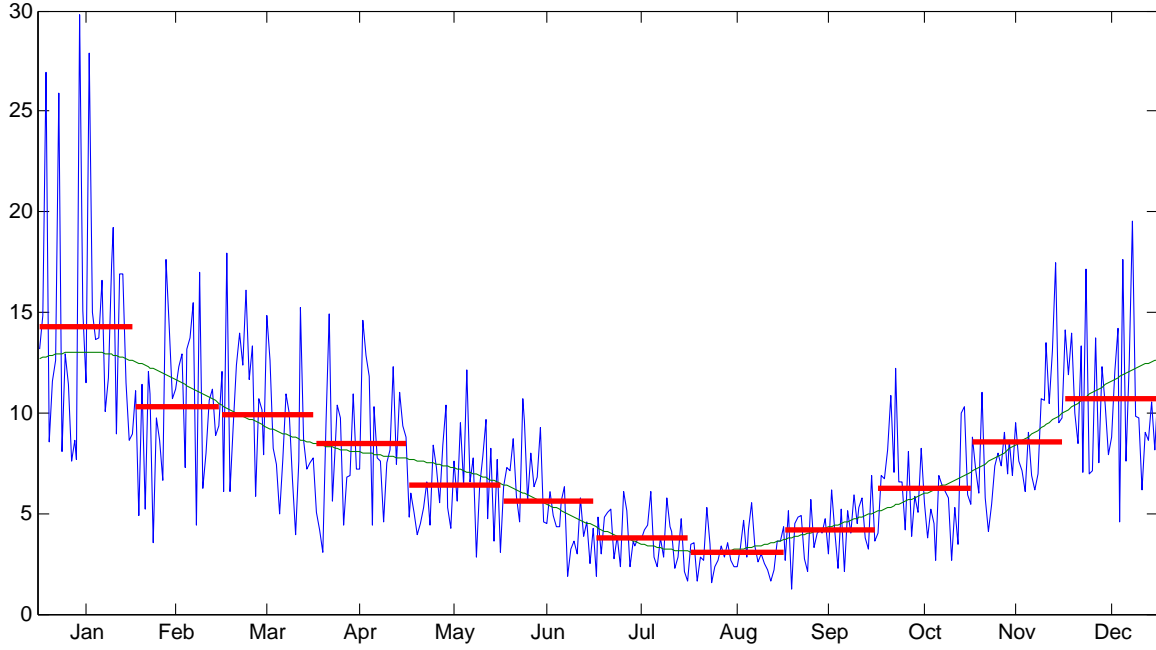


Figure 5: Square residuals $(\epsilon^T)^2$ (blue) for New York La Guardia temperatures, compared with $\sigma_T^2(t)$ fitted with (5) and the monthly average square residuals (red) using the same time window of the HH dataset (March 10, 1994 - December 31, 2014).

temperatures on time interval of one day up to 3 days.

If some dependencies are observed between natural gas prices and temperatures measured in the future, *ex post* we can argue that temperatures at future days could have influenced prices only for the part of information that could be forecasted by the most accurate weather forecasting service. In this case we claim that gas prices include the piece of information coming from temperatures forecast.

In order to have a simple estimation of correlation, we would like to calibrate this model parameter directly on residuals, avoiding considering the hidden random variable z at time t . Next **proposition** states that this is possible for the model presented in this study.

Proposition 1. If residuals are described by equation (6) and their correlation by equation (7), then the correlation parameter can be obtained via the relation

$$\rho(t) = \frac{1}{\sigma_T(t + \tau)} \mathbb{E}[\Delta I(t) \Delta J(t + \tau)] \left(\frac{\delta}{\gamma} \right)^{-1/2} \frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)} \quad (8)$$

with $\gamma = \sqrt{\alpha^2 - \beta^2}$ and $K_\nu(w)$ the Modified Bessel function of the second kind with $\nu, w \geq 0$. *Proof:* See Appendix A.

4 Model calibration and statistical significance

In this section we present the findings from the empirical analysis, discussing in detail the significance of the calibrated parameters.

Estimating parameters' significance could be not trivial using standard statistical methods, even when considering separately the time series of temperature and natural gas. The simplest

approach for achieving this task is to consider the statistical bootstrap method described in Efron and Tibshirani (1986). It is a numerical method very useful to estimate confidence intervals of the calibrated quantities when, as in the case of the model presented in this study, the theoretical analysis is not trivial and no simple expression exists for confidence intervals. We consider separately the gas model and the temperature model: starting from the calibrated set of parameters, we reproduce M time series associating to each date in the series a new value obtained via a Monte Carlo simulation of equations (1) and (3). Each simulated time series (called a bootstrap sample) has the same size of the original time series, i.e. 5203 prices in the natural gas case and 7594 values in the temperatures series for the three cities. We consider $M = 10^4$ and for each bootstrap sample we calibrate again model parameters. The confidence intervals at the ξ confidence level are built valuing for each parameter the $(1 - \xi)/2$ and the $(1 + \xi)/2$ percentiles of the sample; similarly it is also obtained the statistical significance of the parameters at 10%, 5% and 1% significance level.

seasonality	\hat{a}_0	4.307***	(3.095, 5.584)
	\hat{a}_1	0.167	(−0.149, 0.487)
	\hat{a}_2	−0.117	(−0.487, 0.176)
	\hat{a}_3	−0.229***	(−0.410, −0.063)
	\hat{a}_4	−0.085	(−0.271, 0.057)
AR(1)	\hat{b}_1	0.996***	(0.992, 0.998)
NIG	$\hat{\alpha}$	13.999***	(12.606, 15.841)
	$\hat{\beta}$	−0.248***	(−0.264, −0.084)
	$\hat{\delta}$	0.027***	(0.025, 0.029)

Table 1: *Parameters of gas spot model calibrated with HH data from March 10, 1994 to December 31, 2014. Each parameter is followed by its confidence interval at 95% level. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.*

In this section first we briefly summarize calibration and statistical significance for temperatures and natural gas analysed as separate processes (subsection 4.1). We then focus on calibration and statistical significance of the two main joint model parameters and discuss the results achieved (subsections 4.2 and 4.3). Finally we consider the robustness of the analysis (subsection 4.4).

4.1 Calibration and significance for gas and temperature models

In table 1 we show the calibrated parameters for natural gas and in table 2 the calibrated parameters for temperature in the three cities under analysis. Confidence intervals are evaluated at the 95% confidence level; we also report whether parameters' are statistical significant at 10%, 5% and 1% significance level.

We observe that in the seasonality of gas spot \hat{a}_0 and \hat{a}_3 are significant, i.e. the semiannual period in the seasonal mean function is significant. This fact confirms our choice of using (2) with $p = 2$. NIG residuals have been calibrated via a Maximum Likelihood estimation of the 3 parameters (i.e. with $\mu = 0$); this choice is motivated by a not significant μ in the 4 parameters case. We show that such NIG distribution describes accurately the observed residuals' distribution (see figure 6) and NIG parameters are all significant at 1% significance

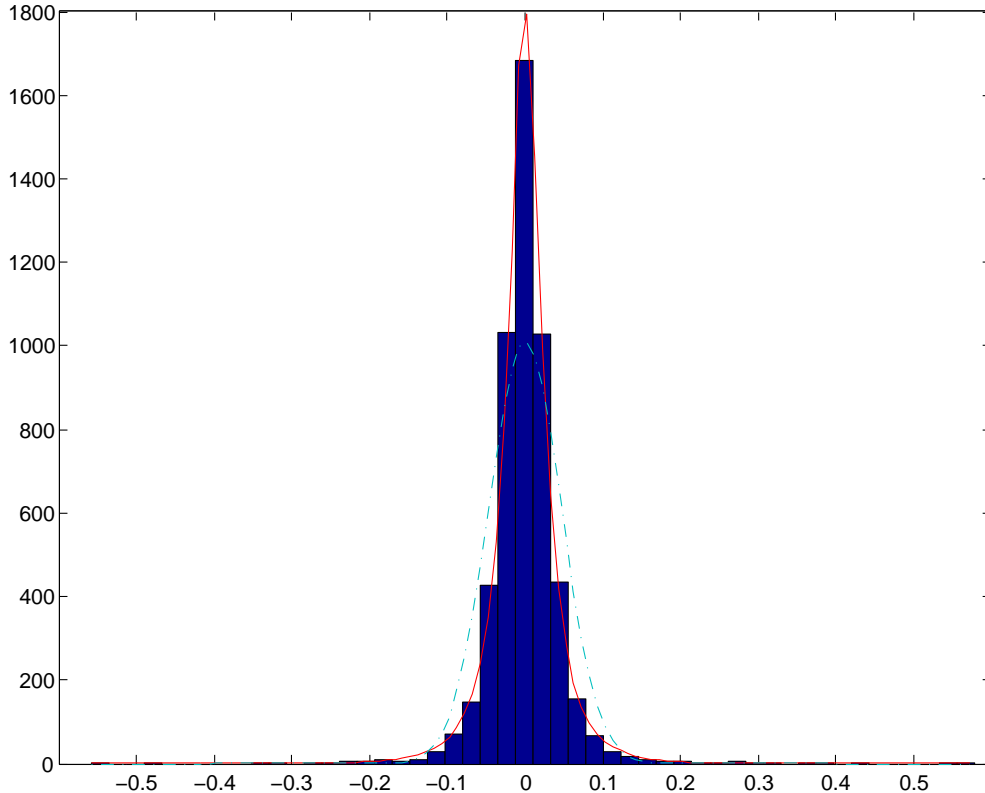


Figure 6: *Plot of the histogram (blue) of natural gas residuals compared with the NIG distribution with the calibrated parameters obtained via the Maximum Likelihood estimation (red with continuous line). We plot also the Gaussian distribution with the same variance (light blue with dot-dashed line), showing the clear presence of fat tails.*

level (see table 1). The same significance holds also for the autoregressive parameter: let us notice that the estimate \hat{b}_1 is close to 1 (i.e. a very slow mean reversion, equivalent to a decay time of approximately 249 days) but significantly different from 1 at the 1% level.

Concerning the temperature model, all the calibrated quantities, except the drift term of the seasonality, are significant for the three cities under analysis. Let us mention that extending the analysis to the whole temperature dataset (i.e. starting from the beginning of 1949) the drift parameter becomes positive and significant at 1% significance level (this analysis is not reported in this paper).

4.2 Significance level of correlation

In order to verify the statistical significance of the correlation measure we build an hypothesis test. We consider the distributional properties of this parameter under the zero-hypothesis (H_0) of independence of temperature and natural gas. Since even the probability distribution of the estimator $\hat{\rho}$ under this zero-hypothesis has no simple expression, we perform a bootstrap method in order to evaluate its probability distribution.

We divide the dataset either in months or in seasons: as in Roll (1984), winter is defined as the period including December, January, February; spring, summer and autumn include, respectively, each subsequent three months. We consider either a monthly or a seasonal correlation estimation of the proposed model: we have about $n_m = 430$ datas each month and about $n_s = 1300$ each season. We simulate $M = 10^4$ bootstrap samples either of size n_m or n_s . In

New York			
seasonality	\hat{a}_0^T	13.201***	(12.816 , 13.553)
	\hat{a}_d^T	$3.20 \cdot 10^{-5}$	$(-4.87 \cdot 10^{-5} , 1.17 \cdot 10^{-4})$
	\hat{a}_1^S	-8.967***	(-9.228 , -8.708)
	\hat{a}_2^S	8.234***	(7.973 , 8.490)
AR(3)	\hat{c}_1	0.818***	(0.793 , 0.842)
	\hat{c}_2	-0.279***	(-0.309 , -0.249)
	\hat{c}_3	0.124***	(0.099 , 0.148)

Boston			
seasonality	\hat{a}_0^T	10.807***	(10.447 , 11.155)
	\hat{a}_d^T	$6.03 \cdot 10^{-5}$	$(-2.09 \cdot 10^{-5} , 1.41 \cdot 10^{-4})$
	\hat{a}_1^S	-8.851***	(-9.104 , -8.596)
	\hat{a}_2^S	8.013***	(7.761 , 8.267)
AR(3)	\hat{c}_1	0.776***	(0.751 , 0.799)
	\hat{c}_2	-0.260***	(-0.290 , -0.230)
	\hat{c}_3	0.114***	(0.090 , 0.138)

Chicago			
seasonality	\hat{a}_0^T	10.364***	(9.851 , 10.869)
	\hat{a}_d^T	$8.72 \cdot 10^{-6}$	$(-1.06 \cdot 10^{-4} , 1.23 \cdot 10^{-4})$
	\hat{a}_1^S	-9.448***	(-9.816 , -9.081)
	\hat{a}_2^S	10.106***	(9.747 , 10.461)
AR(3)	\hat{c}_1	0.882***	(0.856 , 0.905)
	\hat{c}_2	-0.290***	(-0.322 , -0.258)
	\hat{c}_3	0.124***	(0.099 , 0.149)

Table 2: *Parameters of temperature model calibrated with the temperatures of New York, Boston and Chicago, from March 10, 1994 to December 31, 2014. Each parameter is followed by its confidence interval at 95% level. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.*

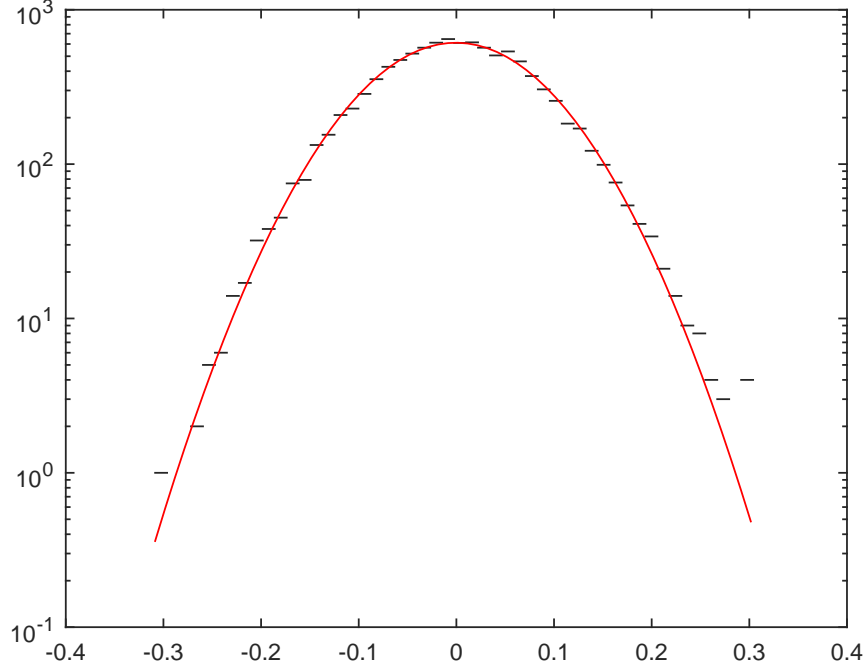


Figure 7: *Log-linear plot of the correlations' frequencies histogram (blue) of $\{\tilde{\rho}_k\}_{k=1,\dots,M=10^4}$. We have simulated M sets of $n = 430$ pairs of reciprocally independent and i.i.d. random variables $\{\tilde{I}_i, \tilde{J}_i\}_{i=1,\dots,n}$, where $\tilde{J}_i \sim N(0, 1)$ and $\tilde{I}_i \sim NIG(\hat{\alpha}, \hat{\beta}, \hat{\delta})$ with $\hat{\alpha}, \hat{\beta}, \hat{\delta}$ values reported in table 1; $\tilde{\rho}_k$ is the correlation measured from the k^{th} set. We have also compared this “experimental” distribution with the one obtained in the standard normal case (red) with standard deviation equal to the σ_ρ indicated in (9).*

Figure 7 we show the histogram of M simulated monthly sequences of $\hat{\rho}$ in the independent case. The distribution of correlation estimator under H_0 hypothesis looks very close to a normal distribution; even through Kolmogorov-Smirnov statistics (DeGroot and Schervish 1986), in both the monthly and seasonal case, we cannot reject the Normality Hypothesis at the 10% significance level.

The standard deviation of $\hat{\rho}$ can be evaluated analytically, as shown in the following **proposition**.

Proposition 2. Let $\{\tilde{I}_i\}_{i=1,\dots,n} \sim NIG(\alpha, \beta, \delta)$ i.i.d. and $\{\tilde{J}_i\}_{i=1,\dots,n} \sim N(0, \sigma^2)$ i.i.d. be two independent samples. Then the standard deviation of the sample correlation coefficient $\hat{\rho}$ is

$$\sigma_\rho = \sqrt{\frac{1}{n} \left[\beta^2 \left(\frac{1}{\gamma^2} + \frac{\delta}{\gamma} \right) + 1 \right] \frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)}} \quad (9)$$

Proof: See Appendix A.

A consequence of this proposition is that we are able to verify the significance of $\hat{\rho}$ just computing the p-value under the H_0 hypothesis.

4.3 Main results and the role of the delay parameter τ

The significance of the parameter $\hat{\rho}$ is also useful in indicating the most relevant values for the delay τ .

In table 3 we show the seasonal $\hat{\rho}$ for the Henry Hub correlated with the temperature of New York, for different values of τ .

	$\tau = 0$		$\tau = 1$		$\tau = 2$		$\tau = 3$	
	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value
Winter	0.0480	16.22%	-0.1998***	(10 ⁻⁸)	-0.2179***	(10 ⁻¹⁰)	-0.1430***	(10 ⁻⁵)
Spring	0.0374	25.64%	-0.0434	18.82%	-0.0817**	1.31%	-0.0320	33.16%
Summer	0.0528	10.60%	0.0471	14.93%	0.0385	23.87%	0.0178	58.48%
Autumn	0.0542	10.47%	-0.1379***	(10 ⁻⁵)	-0.0600*	7.25%	-0.0465	16.44%

Table 3: Correlation parameter $\hat{\rho}(t)$ estimated seasonally, for $\tau = 1, 2$ and 3, using HH gas spot and New York temperature. For each value we report the significance level (p-value) as described in section 4.2. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.

In the New York case using $\tau = 0$, we do not find significant results, so we can conclude that the gas market doesn't react simultaneously to a temperature variation. In the case of $\tau = 1$ we observe a negative and strongly significant $\hat{\rho}$ for winter and autumn. Even for $\tau = 2$ we get significant values but lower than the ones observed for $\tau = 1$ in autumn, where the negative correlation is significant only at 10%; also in spring it is observed a significant correlation at the 5% level. We conclude that the most significant results are observed considering a delay $\tau = 1$. This isn't a surprising result: in fact the delivery of the gas spot is one day delayed, so it seems natural that trader's decisions are mainly based on temperature forecasts for the delivery day, i.e. the following day. Also for Boston and Chicago the most significant correlation are observed for $\tau = 1$; significant results appear up to $\tau = 3$.

	NewYork ($\tau = 1$)		Boston ($\tau = 1$)		Chicago ($\tau = 1$)	
	$\hat{\rho}(t)$	p-value	$\hat{\rho}(t)$	p-value	$\hat{\rho}(t)$	p-value
winter	-0.1998***	(10 ⁻⁸)	-0.1815***	(10 ⁻⁷)	-0.2294***	(10 ⁻¹¹)
spring	-0.0434	18.82%	-0.0193	55.84%	-0.0362	27.22%
summer	0.0471	14.93%	0.0541*	9.79%	0.0540*	9.85%
autumn	-0.1379***	(10 ⁻⁵)	-0.1086***	0.12%	-0.0764**	2.23%

Table 4: Correlation parameter $\hat{\rho}(t)$ estimated seasonally with $\tau = 1$, using HH gas spot and the temperature of New York, Boston and Chicago. For each value we report the significance level (p-value) as described in 4.2. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.

In tables 4 and 5 we show the results for $\hat{\rho}$ for the three cities in Northeast and Midwest, together with their p-values for $\tau = 1$. The parameter of correlation is estimated both seasonally and monthly.

This is the most relevant result of the paper: significant negative correlations can be observed during the cold period. We obtain strongly significant values for the cold period: in particular November, January and February are the most correlated months. Autumn and winter are

	NewYork ($\tau = 1$)		Boston ($\tau = 1$)		Chicago ($\tau = 1$)	
	$\hat{\rho}(t)$	p-value	$\hat{\rho}(t)$	p-value	$\hat{\rho}(t)$	p-value
Jan	-0.3196***	(10^{-7})	-0.2829***	(10^{-6})	-0.2780***	(10^{-6})
Feb	-0.1939***	0.17%	-0.2278***	0.02%	-0.2432***	0.01%
Mar	-0.0762	17.85%	-0.0197	72.77%	-0.0928	10.15%
Apr	-0.0465	41.99%	-0.0185	74.78%	-0.0262	64.97%
May	0.0009	98.77%	-0.0198	72.87%	0.0199	72.73%
Jun	0.0247	66.41%	0.0192	73.57%	0.0342	54.67%
Jul	0.0205	71.97%	0.0301	59.88%	0.1180**	3.89%
Aug	0.1054*	5.89%	0.1217**	2.91%	0.0126	82.15%
Sep	0.0016	97.79%	-0.0494	39.86%	0.0531	36.45%
Oct	-0.1015*	6.95%	-0.0618	26.93%	-0.0294	59.96%
Nov	-0.2776***	(10^{-6})	-0.1983***	0.08%	-0.2110***	0.04%
Dec	-0.0758	18.87%	-0.0348	54.64%	-0.1693***	0.33%

Table 5: Correlation parameter $\hat{\rho}(t)$ estimated monthly for $\tau = 1$, using HH gas spot and the temperature of New York, Boston and Chicago. For each value we report the significance level (p-value) as described in 4.2. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.

always negative and significant (always at 1% level with the exception of Chicago at 5% in autumn), summer is positive and significant at 10% for Chicago and Boston.

4.4 Robustness of the results

Results' robustness has been verified dividing the sample into four subsamples:

1. 10 March 1994 up to 31 Dec 1999 (> 5 years and 9 months);
2. 1 Jan 2000 up to 31 Dec 2003 (4 years);
3. 1 Jan 2004 up to 31 Dec 2008 (5 years);
4. 1 Jan 2009 up to 31 Dec 2014 (6 years).

This division corresponds to some major changes in U.S. natural gas market and to a division of the time series into four non-overlapping samples with the same order of magnitude of data points.

After a first period, between 1994 and 1999, of full liberalization of natural gas market that has been described in Section 2, the first major change corresponds to a boom in construction of new gas-fired electric plants at the beginning of the millennium, largely driven by independent power producers in response to deregulation in the electric power sector. From 2000 through 2003, electric generating capacity representing NG-fired plants increased of an averaged 45 GW annually in the U.S.A.; this value drops to an yearly average lower than 10 GW afterwards (Jones and Martin 2014): a second period focuses on this four year interval starting on the 1st of January 2000 and ending on the 31st of December 2003.

A third period, between the 1st of January 2004 and the end of 2008, is characterized by a significant increase in NG consumption and by higher prices. Winter demand in 2002-2003 was the highest of the previous 5 years, storage inventories declined from 5y heights in October 2002 to 5y lows in March 2003, depleting storage by a record 80%; by early April, storage had

		1994 - 1999		2000 - 2003	
seasonality	\hat{a}_0	2.164***	(1.955; 2.399)	4.254***	(3.446; 5.296)
	\hat{a}_1	0.001	(−0.207; 0.215)	0.384	(−0.358; 1.003)
	a_2	−0.149	(−0.377; 0.079)	0.065	(−0.780; 0.546)
	\hat{a}_3	−0.119*	(−0.257; 0.013)	0.276	(−0.119; 0.636)
	\hat{a}_4	−0.008	(−0.142; 0.132)	−0.274**	(−0.800; −0.013)
AR(1)	\hat{b}_1	0.978***	(0.958; 0.984)	0.986***	(0.964; 0.991)
NIG	$\hat{\alpha}$	9.773***	(7.903; 12.844)	14.753***	(11.876; 19.841)
	$\hat{\beta}$	−0.249***	(−0.408; −0.120)	0.088	(−0.173; 0.312)
	$\hat{\delta}$	0.021***	(0.019; 0.023)	0.035***	(0.031; 0.042)

		2004 - 2008		2009 - 2014	
seasonality	\hat{a}_0	7.424***	(6.336; 8.445)	3.844***	(3.439; 4.236)
	\hat{a}_1	−0.065	(−0.968; 0.733)	0.118	(−0.181; 0.496)
	\hat{a}_2	0.025	(−0.874; 0.812)	0.161	(−0.135; 0.533)
	\hat{a}_3	0.144	(−0.340; 0.603)	0.212**	(0.022; 0.420)
	\hat{a}_4	−0.395*	(−0.927; 0.068)	0.018	(−0.153; 0.251)
AR(1)	\hat{b}_1	0.984***	(0.962; 0.989)	0.982***	(0.964; 0.988)
NIG	$\hat{\alpha}$	21.471***	(17.559; 27.739)	16.792***	(13.926; 21.303)
	$\hat{\beta}$	−0.314*	(−0.367; 0.034)	−0.476***	(−0.562; −0.198)
	$\hat{\delta}$	0.040***	(0.034; 0.048)	0.024***	(0.022; 0.028)

Table 6: *Parameters of gas spot model calibrated with HH data splitted in the four subsamples. Each parameter is followed by its confidence interval at 95% level. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.*

bottomed out below 650 million MMBtu, more than 40% below normal storage for that time of year (Choo 2003). Spot HH prices reached for the first time a price above USD 15 per MMBtu (on the 25th of February 2003) and remained above USD 10 for some days. These spikes signalled industry's limits in physical capacity and storage inventories: HH prices remained higher than USD 5/MMBtu, apart from a few days, for the whole period 2004-2008. The average price in the period was USD 7.42, compared with USD 2.18 in the first period and with USD 4.27 in the second one. One of the first effects of the soaring NG prices was an increase of LNG deliveries in 2003 that has more than doubled with respect to previous year (Mastrangelo 2007).

Finally a fourth period of six years (2009 up to 2014) is characterized by major technological advancements that made exploration more cost effective and by a significant increase of investments in transportation: both drivers play an important role in NG price dynamics. On one hand, the main technological advancement was the use of horizontal drilling technique together with hydraulic fracturing methods in the Marcellus shale gas basin (the largest in U.S.A.), techniques that had previously been applied in the Barnett shale of Texas; the high yield wells discovered at the beginning of 2008 have showed a cost lower than expected and then they have proved much more efficient and productive (Engelder 2009). The percentage of U.S. natural gas production from shales has increased from the 10% of total dry natural gas production in U.S.A. at the beginning of 2008 up to the 39% at the end of 2012 (EIA 2013b). On the other hand, transmission and distribution network is also important; during 2008 the completions of new NG pipelines and expansions of the existing ones represented the greatest amount of

pipeline activity in more than 5 years in U.S.A.: the increase was three times the previous 5y average in terms of both mileage and capacity (Gaul 2009)¹. During this fourth period prices have lowered to an average of USD 4.57 per MMBtu and to USD 3.61 just considering the last three years.

We observe that, even if the four selected periods look very different in the historical perspective as we have summarized above, the same *stylized facts* appear in the natural gas market in all four periods.

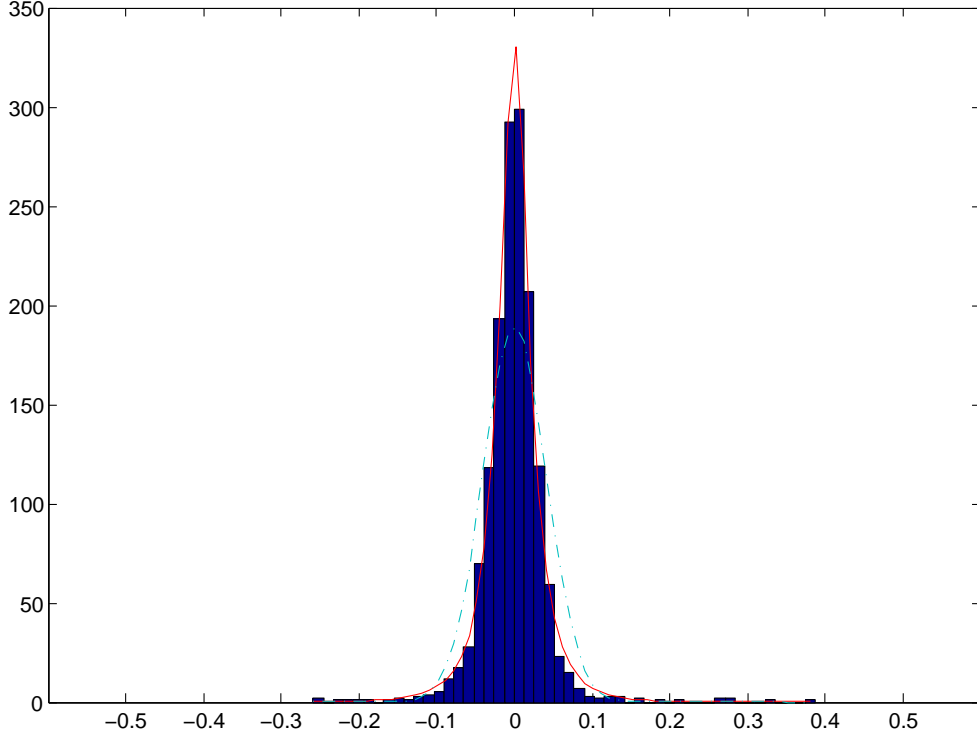


Figure 8: *Plot of the histogram (blue) of natural gas residuals in the 4th period compared with the NIG distribution with the calibrated parameters obtained via the Maximum Likelihood estimation (red with continuous line). We plot also the Gaussian distribution with the same variance of the sample (light blue with dot-dashed line). The shape of the distribution is qualitatively similar to the one for the whole time-series of figure 6, showing even in this case the clear presence of fat tails.*

First, considering the statistical properties of NG time-series separately from the temperatures, we notice that the main characteristics are confirmed in each period.

As shown in table 6 the NG time-series are always well described by an AR(1) process with NIG residuals with all parameters always significant at 1% (with the only exception of β significant at 10% in the 3rd period and not significant in the 2nd period, signalling a non-skewed distribution during this four-year period), where confidence intervals are computed with the bootstrap technique described at the beginning of Section 4. In figure 8 we show that NIG distribution describes accurately the observed residuals' distribution even focusing just on the last period

¹A significant example of this activity is the Millennium Pipeline: a newly constructed hi-capacity NG transmission line, that opened on December 22, 2008, transporting gas from the heart of the Marcellus shale region along the Pennsylvania border to New York City and its metropolitan area.

(2009-2014): residuals' distribution looks very similar to the one obtained considering the whole sample of 20 years showed in figure 6. Similar results hold for the other three periods.

Regarding the seasonality, in table 6 we also observe that the semiannual seasonality coefficient is always significant (at 5% confidence level in the 2nd and the 4th periods, at 10% level in the 1st and the 3rd periods). For example in figure 9 we plot the average daily HH price and its seasonality in the last period: it is possible to recognize clearly the two peaks in winter and in summer of the semi-annual seasonality; even in this case, as in figure 3, the yearly seasonality (dashed line) does not look adequate.

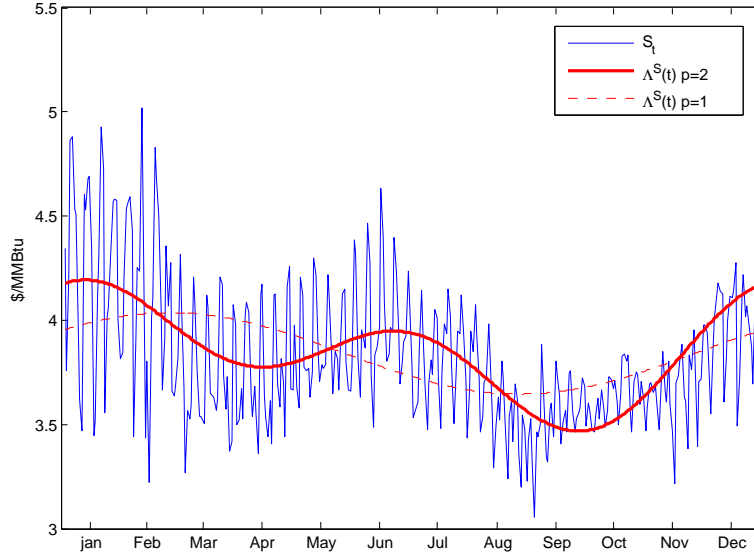


Figure 9: *Average daily Henry Hub gas price (blue line) in the 4th period from January 1, 2009 to December 31, 2014 and seasonality $\Lambda^S(t)$ (red) in \$/MMBtu. We observe that the yearly seasonality (dashed line) does not describe properly the data, as in the whole sample (see e.g. figure 3).*

Second, the observed pattern for correlation between HH prices and temperatures during the colder season is confirmed in each one of the four periods. Estimated correlations and their p-values for the four seasons in the four periods are shown in Table 7, while in Appendix B we report monthly results in the New York case. In winter, correlation is always negative and significant for the three cities (either at 1% or at the 5% level), in autumn, correlation is always negative and it is significant in 6 cases out of 12 (at different levels).

It is interesting to notice that during the third period, a positive correlation in the summer season is observed, significant for New York and Boston at the 5% confidence level: this period (i.e. between 2004 and 2008) has followed the boom in new gas-fired electric generation and it was characterized by a growing natural gas demand in the summer season.

Let us summarize the robustness analysis realized: we have divided the sample in four periods characterized by major differences in the NG market and we can confirm the main results of this study. In particular we observe a seasonal pattern for one of the main parameters in the joint model: gas-temperature correlation, negative in the cold season, becomes positive in the warmer part of the year (and statistically significant only during the third period of this analysis for Boston and New York).

	NewYork('94-'99)		NewYork('00-'03)		NewYork('04-'08)		NewYork('09-'14)	
	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value
Win	-0.172**	1.78%	-0.318***	10^{-5}	-0.213***	0.10%	-0.145**	2.08%
Spr	-0.041	54.75%	-0.019	79.74%	-0.021	73.58%	-0.119*	5.08%
Sum	-0.011	87.44%	0.013	85.46%	0.129**	3.72%	0.050	40.42%
Aut	-0.219***	0.12%	-0.163**	2.69%	-0.059	36.19%	-0.122**	4.73%

	Boston('94-'99)		Boston('00-'03)		Boston('04-'08)		Boston('09-'14)	
	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value
Win	-0.163**	2.47%	-0.318***	10^{-5}	-0.145**	2.43%	-0.145**	2.11%
Spr	0.006	93.13%	0.013	85.91%	-0.018	77.31%	-0.095	11.74%
Sum	-0.003	96.28%	0.035	62.83%	0.140**	2.32%	0.043	47.56%
Aut	-0.186***	0.59%	-0.152**	3.86%	-0.027	67.65%	-0.086	16.46%

	Chicago('94-'99)		Chicago('00-'03)		Chicago('04-'08)		Chicago('09-'14)	
	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value
Win	-0.449***	10^{-9}	-0.152**	4.36%	-0.182***	0.48%	-0.169***	0.72%
Spr	-0.133**	4.96%	0.012	86.65%	-0.049	43.44%	-0.012	83.92%
Sum	0.058	38.60%	-0.009	89.80%	0.066	28.83%	0.092	12.84%
Aut	-0.113*	9.63%	-0.071	33.68%	-0.065	30.82%	-0.066	28.52%

Table 7: Correlation parameter $\hat{\rho}(t)$ seasonally with $\tau = 1$, using HH gas spot data splitted in the four subsamples and the temperature of New York, Boston and Chicago. For each value we report the significance level (p-value) as described in 4.2. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.

As a further check of robustness we also consider the case where we estimate the correlation between every business day gas price and the average of the temperatures on the next business day and non-business days if there are any before the next business day; the main difference is on Friday HH prices that are related to the average of next Saturday, Sunday and Monday temperatures. Results are quite similar with regard to seasonal and monthly correlations: as we expect the most reliable temperature forecast (i.e. the one relative to the next calendar day) plays the most relevant role. Boston shows similar values, while both Chicago and New York exhibit more significant results. For Chicago autumn correlation becomes significant at 1% level (with summer not anymore significant at 10%) and March correlation is significant at 10% level, while for New York spring and March correlations become significant at 10% level and August (with a positive correlation) and December significant at 5% level.

5 Conclusions

In this study we have proposed a joint model for temperatures and natural gas; we have also shown that future temperatures have an impact on natural gas prices observing that this market is subject to demand shocks related to temperature.

A simple description for natural gas and temperature looks accurate. Both dynamics share the same structure with a seasonality and an AR component; moreover, residuals are NIG for natural gas and they are Gaussian for temperatures. Their dynamics is joined via a correlation parameter and a time-delay. With such an elementary model we can provide statistical

significance of all calibrated parameters.

Despite model simplicity we are able to obtain significant results in the U.S. market for the main model parameters. We have analysed the complete time-series of Henry Hub spot prices in the period between the 10th of March 1994 and the end of 2014 and the corresponding temperatures of three cities in Northeast and Midwest regions of U.S.A. We show during colder periods a negative correlation between Henry Hub gas spot prices and the temperatures measured the next day in the three cities. The negative correlation observed is statistically significant and in most cases at 1% level: a relevant (and not obvious) result for the liquid market of natural gas. The results are robust even dividing the sample into four time intervals characterized by major differences in the natural gas market; we also observe a significant positive correlation in the 2004-2008 period that has followed a remarkable increase in new NG power generation.

Let us mention some possible applications of the model presented in this paper.

Implications of this model for natural gas market makers are rather obvious since next days' temperature forecasts provide relevant information in price formation in the day-ahead market. Moreover, a joint model of gas price and temperature can be used for pricing and risk management purposes. A possible example is a derivative contract known as "temperature swap", where the payoff is a function of both the temperature in some specific locations (above a given temperature strike) and the NG price (above a given strike). Another typical example is a retail contract price by a gas supplier taking into account also volume risk: customers often pay a fixed price (per unit of gas delivered) for their consumption with no constraints on volumes; since there is a known deterministic dependency between gas volumes and temperature, a joint model of gas price and temperature allows taking into account volume-related risks.

Acknowledgements

We thank all participants to the Second Conference on Stochastics of Environmental and Financial Economics at CAS in Oslo 2015 and to the Energy Finance Conference in London 2015 and in particular M. Coulon, A. Müller and R. Stelzer for useful comments. We feel indebted to P. Laurence for several helpful and wise suggestions on the subject and R. Carmona for an enlightening discussion on this topic at Vienna's Wolfgang Pauli Institute. We thank also M. Brunetti and C. Simolo of the Institute of Atmospheric Sciences and Climate (CNR), J. Robel of the National Climatic Data Center and C. Andrews of EIA for some help on temperatures and natural gas datasets. The usual disclaimers apply.

References

- Barndorff-Nielsen, O.E., 1997. Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling, *Scandinavian Journal of statistics*, 24, 1–13.
- Benth, F.E. and Benth, J.S., 2004. The normal inverse Gaussian distribution and spot price modelling in energy markets, *International Journal of Theoretical and Applied Finance*, 7, 177–192.
- Benth, F.E. and Benth, J.S., 2007. The volatility of temperature and pricing of weather derivatives, *Quantitative Finance*, 7, 553–561.
- Benth, F.E. and Benth, J.S., 2012. A critical view on temperature modelling for application in weather derivatives markets, *Energy Economics*, 34, 592–602.

- Benth, F.E., Benth, J.S., and Koekebakker, S., 2008. *Stochastic modelling of electricity and related markets*, World Scientific Publishing Co. Pte. Ltd.
- Campbell, S.D. and Diebold, F.X., 2005. Weather forecasting for weather derivatives, *Journal of the American Statistical Association*, 100, 6–16.
- Choo, L.F., 2003. Report of the natural gas spike of february 2003, *Federal Energy Regulatory Commission (July 13)*, 1–16.
- DeGroot, M.H. and Schervish, M.J., 1986. *Probability and statistics*, Addison-Wesley Reading, MA, 4th ed.
- Efron, B. and Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statistical science*, 1, 54–75.
- Engelder, T., 2009. Marcellus 2008: Report card on the breakout year for gas production in the Appalachian basin, *Fort Worth Basin Oil and Gas Magazine*, 7, 18–22.
- Gaul, D., 2009. Expansion of the U.S. natural gas pipeline network: Additions in 2008 and projects through 2011, *Office of Oil and Gas, EIA Report*, 1–17.
- Gunnarshaug, J. and Ellerman, A.D., 1998. Natural gas pricing in the Northeastern U.S., *MIT Center for Energy and Environmental Policy Research*.
- Jones, J. and Martin, L., 2014. EIA projects modest needs for new electric generation capacity, *Today in Energy*.
- Lott, N., Ross, T., and Sittel, M., 1996. The winter of 95-96. A season of extremes, *National Climatic Data Center Technical Report (April)*, 1–32.
- Mastrangelo, E., 2007. Natural gas year in review 2006, *Office of Oil and Gas, EIA Report*, 1–6.
- Müller, J., Hirsch, G., and Müller, A., 2015. Modeling the price of natural gas with temperature and oil price as exogenous factors, in: *Innovations in Quantitative Risk Management*, Springer International Publishing, 109–128.
- Preston, J., 2014. 'Polar vortex' brings coldest temperatures in decades, *The Lede - The New York Times (January 6)*.
- Roll, R., 1984. Orange juice and weather, *The American Economic Review*, 74, 861–880.
- Schwartz, E.S., 1997. The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging, *The Journal of Finance*, 52, 923–973.
- Stoll, S.O. and Wiebauer, K., 2010. A spot price model for natural gas considering temperature as an exogenous factor with application, *The Journal of Energy Markets*, 3, 113–128.
- Weron, R., 2014. Electricity price forecasting: A review of the state-of-the-art with a look into the future, *International Journal of Forecasting*, 30, 1030–1081.
- U.S. Energy Information Administration, 2013a. Heating and cooling no longer majority of U.S. home energy use, *Today in Energy*.

U.S. Energy Information Administration, 2013b. Natural gas weekly update (December, 19), *Independent Statistics & Analysis*.

U.S. Energy Information Administration, 2014. Factors affecting natural gas prices, *Natural Gas Explained*.

U.S. Energy Information Administration, 2015. Natural gas monthly (January), *Independent Statistics & Analysis*, 1–123.

Appendix A

Proof of **proposition 1** is a consequence of the following **lemma**.

Lemma: Let $I \sim NIG(\alpha, \beta, \delta)$, i.e.

$$I = \beta z + \sqrt{z}g$$

with z an IG mixing variable and g a standard normal variable. Let $J \sim N(0, \sigma^2)$, a random variable independent from z , s.t. $\mathbb{E}[gJ] = \rho\sigma$. Then

$$\mathbb{E}[IJ] = \rho\sigma \left(\frac{\delta}{\gamma}\right)^{1/2} \frac{K_0(\delta\gamma)}{K_{1/2}(\delta\gamma)}.$$

Proof: Straightforward by direct inspection

$$\mathbb{E}[IJ] = \rho\sigma\mathbb{E}[z^{1/2}] = \rho\sigma \left(\frac{\delta}{\gamma}\right)^{1/2} \frac{K_0(\delta\gamma)}{K_{1/2}(\delta\gamma)}$$

using the formula for IG moments, or equivalently

$$\rho = \mathbb{E}[IJ] \left(\frac{\gamma}{\delta\sigma^2}\right)^{1/2} \frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)} \quad \clubsuit \quad (10)$$

Proof of **proposition 2:** Sample correlation coefficient of equation (10) is

$$\hat{\rho} = \frac{1}{n} \left(\sum_{i=1}^n \tilde{I}_i \tilde{J}_i \right) \left(\frac{\gamma}{\delta\sigma^2} \right)^{1/2} \frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)}.$$

Given the independence of the two sets of random variables, it is straightforward to show that $\mathbb{E}[\hat{\rho}] = 0$. Then $\hat{\rho}$ variance is

$$\begin{aligned} \sigma_{\hat{\rho}}^2 &= \mathbb{E} \left[\left(\frac{1}{n} \left(\sum_{i=1}^n \tilde{I}_i \tilde{J}_i \right) \left(\frac{\gamma}{\delta\sigma^2} \right)^{1/2} \frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)} \right)^2 \right] \\ &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{E} \left[\left(\tilde{I}_i \tilde{J}_i \right)^2 \right] \frac{\gamma}{\delta\sigma^2} \left(\frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)} \right)^2 = \\ &= \frac{1}{n} \mathbb{E} \left[\left(\tilde{I}_1 \tilde{J}_1 \right)^2 \right] \frac{\gamma}{\delta\sigma^2} \left(\frac{K_{1/2}(\delta\gamma)}{K_0(\delta\gamma)} \right)^2 \end{aligned}$$

and

$$\mathbb{E}[(\tilde{I}_1 \tilde{J}_1)^2] = \mathbb{E}[\tilde{I}_1^2] \mathbb{E}[\tilde{J}_1^2] = \frac{\delta\sigma^2}{\gamma} \left[\beta^2 \left(\frac{1}{\gamma^2} + \frac{\delta}{\gamma} \right) + 1 \right] \quad \clubsuit$$

Appendix B

In this appendix we show some other outcomes on the robustness analysis of the results.

We report monthly robustness results for New-York in table 8. We observe the same qualitative pattern discussed for seasonal results in Section 4.4: colder months present negative correlation (significant in all 4 periods during January and in 3 periods out of 4 during November) while August presents a positive and significant correlation in the 3rd period at 5%. Clearly the lower significance of the results is related to the smaller sample sizes examined.

	NewYork('94-'99)		NewYork('00-'03)		NewYork('04-'08)		NewYork('09-'14)	
	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value	$\rho(t)$	p-value
Jan	−0.4385***	0.06%	−0.3063**	1.68%	−0.3195***	0.41%	−0.1872*	8.46%
Feb	0.0203	88.00%	−0.5002***	0.02%	−0.1863	10.15%	−0.1656	13.71%
Mar	0.0001	99.96%	−0.0634	61.21%	−0.0940	37.41%	−0.1720*	9.78%
Apr	−0.1886	10.72%	0.1052	41.14%	0.0163	88.13%	−0.1193	26.19%
May	0.0782	50.25%	−0.0955	44.24%	0.0335	75.55%	−0.0515	62.70%
Jun	0.0545	63.62%	−0.0281	82.44%	0.0609	56.96%	0.0011	99.18%
Jul	−0.0144	90.15%	−0.0444	72.56%	0.0891	41.46%	0.0271	79.68%
Aug	−0.0876	44.19%	0.1372	26.45%	0.2497**	1.76%	0.1459	15.88%
Sep	0.0546	64.21%	0.0316	80.73%	−0.0027	98.11%	−0.0675	52.93%
Oct	−0.2196*	5.39%	−0.1299	28.53%	0.0056	95.87%	−0.0722	48.41%
Nov	−0.4187***	0.05%	−0.3760***	0.44%	−0.1462	19.18%	−0.2157*	5.08%
Dec	−0.0318	78.42%	−0.1502	24.95%	−0.1255	25.51%	−0.0823	43.90%

Table 8: *Correlation parameter $\hat{\rho}(t)$ estimated monthly for $\tau = 1$, using HH gas spot data splitted in the four subsamples and the temperature of New York. For each value we report the significance level (p-value) as described in 4.2. With *, **, *** we indicate statistical significance of the parameters at 10%, 5% and 1% significance level, respectively.*

As another test of robustness we also consider four non-overlapping samples with the same number of data points. All results are confirmed.

Finally we also plot NG correlation for the three most relevant seasons in this study (winter, summer and autumn) with 5y overlapping windows for the city of New York. This plot provides just a qualitative indication of time evolution for correlation due to the obvious statistical implication of considering overlapping intervals. We indicate in the horizontal line of the plot the last year of the 5y interval (e.g. 1999 indicates the 1995-1999 period).

As one could expect winter correlations become less pronounced after 2008 due to shale gas production (that has grown especially in the Marcellus basin, geographically near to the most important NG-demand centers in U.S.A.) and better elasticity of supply; we observe that winter values are partially compensated by lower values for correlation in autumn. Furthermore, these results suggest also an interesting pattern in the behaviour of correlation, with significant negative correlation observed in the periods of greater consumption that include the two coldest winters in Midwest/Northeast in the time interval analysed, i.e. 1996 and 2014 (Lott *et al.* 1996, Preston 2014) and with significant positive correlation during the summer season in the period that includes (or follows immediately) the years 2000-2003 with an increase in new NG-power plants (2006-2010 in the figure)².

²We thank one referee for having suggested this analysis.

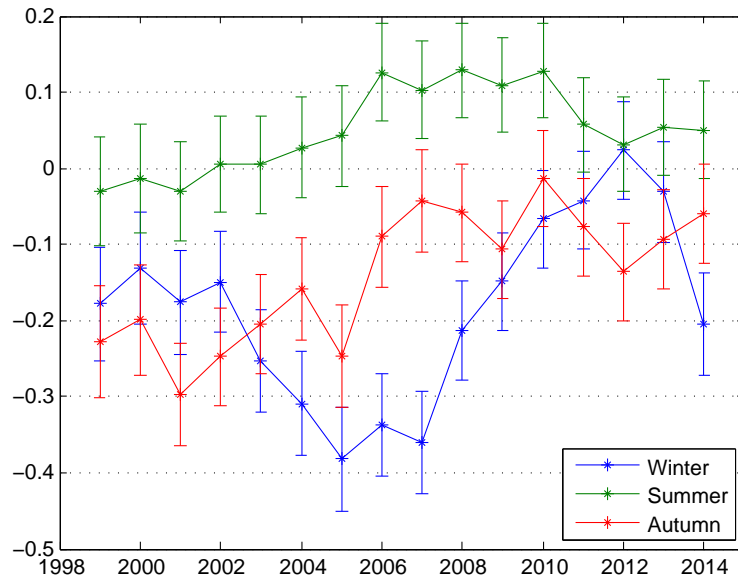


Figure 10: *Correlation estimated with 5y overlapping windows for New York La Guardia. We report the observed correlation with an error bar with \pm one standard deviation for the three seasons most relevant in this study (winter in blue, summer in green and autumn in red).*