THESE DE DOCTORAT EN ECONOMIE

L'UNIVERSITE DE NANTES

ECOLE DOCTORALE N° 597
Sciences Economiques et sciences De Gestion
Spécialité : « *Science économique*»

Par
« Arthur THOMAS »

« The Econometrics of Energy Demand: Identification and Forecast »

Thèse présentée et soutenue à IFP School (ENSPM- Ecole nationale supérieure du pétrole et des moteurs), le 14/12/2020
Unité de recherche : IFPEN/IFP School et LEMNA

Rapporteurs avant soutenance :

Karim Abadir Emeritus Professor of Financial Econometrics, Imperial College London
Dimitris Korobilis Professor of Econometrics, University of Glasgow

Composition du Jury :

Président :	Valérie Mignon	Professeur des Universités, Université Paris Nanterre
Examinateurs :	Karim Abadir	Professor of Financial Econometrics, Imperial College London
	Derek Bunn	Professor of Decision Sciences, Management Science and Operations London Business School
	Dimitris Korobilis	Professor of Econometrics, University of Glasgow
Dir. de thèse :	Benoît Sévi	Professeur des Universités, Université de Nantes
Co-dir. de thèse :	Olivier Massol	Associate Professor in Economics, IFP School and City, University of London

Acknowledgements

First of all, I would like to express my deep and sincere gratitude to my supervisors Prof. Benoit Sévi and Prof. Olivier Massol. Thank you for your continuous support, your patience, motivation, and immense knowledge. I would like to express my deep appreciation and a very special thanks to Moussa Zakaria my research partner who had become a close friend, advise me in my doctoral research, future career choices, and for being part of the people who gave me the joy in the pursuit of knowledge and to awaken my taste for academic research. I would like to thank my thesis committee: Prof. Abadir Karim, Prof. Bunn Derek, Prof. Korobilis Dimitris and Prof. Mignon Valérie for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives. Next, I take this opportunity to express gratitude to Anna Creti, Yannick Le Pen, Giles de Truchis, Elena Dumitrescu, Stéphane Auray, Emmanuel Hache and Benoît Cheze for their advice, discussions about research and encouragement. Warm thoughts to my colleagues and friends: Benoit, Emmanuel, Pierre C., Gondia, Val, Cyprien, Florian, Pierre G., Benjamin, Jérôme, and others. Finally, many heartfelt thanks to Chantal for your help with my administrative tasks. Very specials thanks to Côme, Clément, Margaux, Anthony and Antoine for the great moments together and the long discussions about everyday life, and to become close friends. I would like also to thank my friend and my research partner Frédéric. Many thanks also for Faris, and Cyrille, my two internship supervisors which without them, I probably wouldn't get a PhD in energy economics.

Also, and despite the distance that separates us, I would like to sincerely thank my family: my parents Philippe and Florence, my brothers, Hugo and Maxent, and my sister Romane.

Thank you, grandmothers, Marie and Kasimiria and grandfathers, Bruno and Jean to be proud of me at every moment. Special thanks to Hugo and my father Philippe, for showing me that happiness is different from success and money, and for their consideration about life and the way to live it.

I will thank my friend, Thibaut, Pierrot, Jean-François, Nico, Marie, Florian and Thomas, for the moments I spent with them, trying to forget my work.

Last but not least, I would like now to state my gratitude to Eline even if words are short to express my thankfulness to her. Thank you for your patience, constant support and motivation with me, particularly during the preparation of the defence and uncontrollable fits of laughter. Thank you for your big heart and your capacity to forgive, and for showing me that life is more than work. These years would not have been the same without your love, and I definitely would not have achieved this work, it's also a bit of yours. Thanks to Mimi, for the attacks.

Résumé

La prévention du changement climatique est l'une des priorités de la politique énergétique mondiale qui vise à réduire massivement les émissions de gaz à effet de serre. Face à ces défis, il est frappant de constater que notre connaissance de la modélisation de la demande énergétique demeure imparfaite car elle repose en grande partie sur des travaux empiriques anciens et des méthodologies aujourd'hui dépassées. L'objectif scientifique de cette thèse est double : analyser quantitativement les déterminants économiques de la demande énergétique et développer de nouveaux modèles de prévision. Cette thèse est structurée en quatre chapitres. Le premier chapitre montre que la consommation de gaz naturel en France peut être prédite à l'aide d'un modèle simple utilisant seulement les informations disponibles pour les acteurs du marché. Ce chapitre prouve l'existence d'une relation à long terme entre la demande de gaz naturel et les prix des autres énergies et il fournit des estimations de leurs impacts marginaux sur les niveaux de demande observés. Le deuxième chapitre étudie empiriquement le rôle de la température dans la prévision des prix du gaz aux États-Unis. Il développe une méthodologie de construction d'un nouvel indice mensuel basé sur la température. Cet indice capture les variations de la demande résiduelle de gaz naturel en temps réel. Il est utilisé comme variable exogène supplémentaire dans des modèles structurels VAR afin d'améliorer les prévisions ; et nous montrons que ces modèles prédictifs dérivés de modèles structurels sont améliorés en s'appuyant sur des données en temps réelles (non sujettes à révision). Le troisième chapitre propose d'utiliser dans le cas du pétrole, un modèle structurel capturant les anticipations à l'aide de VAR non causaux et d'identifier correctement les réactions des variables clés du pétrole à un choc d'actualité. Le quatrième chapitre réexamine le pouvoir prédictif de la structure par terme des prix, dite « convenience yield », du pétrole et du gaz en intégrant les anticipations dans une spécification empirique, par le biais d'un VAR non causal basé sur la théorie du stockage qui fournit des prévisions de prix très compétitives dans un cadre bivarié simple.

Mots clés : Demande Energétique ; Prévision ; Identification ; VAR Bayésien ; Non causalité

Abstract

The prevention of climate change is one of the priorities of the world energy policy that aims to massively reduce greenhouse gas emissions. Faced with these challenges, it is striking to note that our knowledge of energy demand modeling remains limited because it is largely based on old empirical work and methodologies that are now dated. Therefore, the objective of our work is twofold. First, we analyze quantitatively the economic determinants of energy demand. Second, we develop new forecasting models. This thesis is structured in four chapters. The first chapter shows that natural gas consumption in France can be predicted using a simple model which only includes public information that is available to market's participants. This chapter proves the existence of a long-term relationship between demand and prices of other energies and provides estimates of their marginal impacts on observed demand levels. The second chapter empirically investigates the role of temperature in forecasting gas prices in the US. It develops a methodology to build a new monthly index based on temperature. This index captures variations in residual demand for natural gas in real time. It is used as an additional exogenous variable in structural models (VAR) to improve forecasts and we show that, in our case, predictive models derived from a structural model are enhanced relying on true real-time (not subject to revisions) data. The third chapter proposes to use, in the case of oil market, a structural model capturing expectations in a non-causal VAR framework, and to properly identify the reactions of oil key variables to supply news shock. The fourth chapter revisits the predictive power of oil and gas convenience yield by incorporating expectations into an empirical specification through non-causal VAR based on the theory of storage which delivers very competitive price predictions in a simple bivariate setting.

Keywords: Energy Demand ; Forecast ; Identification ; Bayesian VAR ; Non causality

Contents

List of Figures

List of Tables

Introduction

Climate change is one of the greatest challenges we face - both in terms of its potential impact on our societies and the planet, and in terms of the extent of international coordination and cooperation that will be required to address it. The scientific evidence is overwhelming. The second article of the Paris Agreement calls for global warming to be kept well below 2°C above pre-industrial levels by 2100 and for continued efforts to limit the rise in temperature to 1.5°C. Meeting these objectives will not be easy, as our societies are energy-dependent at all levels. Energy consumption was necessary, although not sufficient, condition of modern growth (Malanima, 2012). Technical solutions are available, based on different energy sources, but change can take time, due to multiple factors. The inertia of our societies, the slowness of our innovation processes, and the high infrastructure costs. This transition is a transition from the world of carbon energies (Coal, Oil, and Gas) to the so-called renewable energy sources (Wind, Sun, and Biomass), more respectful of the environment. The energy that is gaining consensus in academia to achieve this transition is natural gas (WEO, 2018). It has many advantages (low carbon cost, great flexibility). That's why, our work is more focused on natural gas than oil, but like any transition, to be successful, it requires that we can anticipate and at least partially predict. Indeed, research in this field is experiencing a revival mainly due to two factors. Forecasting models are increasingly powerful and data are increasingly available in terms of volume and frequency. Nevertheless, it must be kept in mind that despite this facility, we should not be tempted to put forecasting before understanding at the risk of becoming very conservative and blind to changes that have not occurred in the past Krivine and Ameisen (2018). For this, it is important to have a well understanding of the path, the current state, and tools that are our disposal. Our work focuses on the implementation of energy demand and prices forecasting models, intending to anticipate and accompany this energy transition. This introduction briefly introduces the history of energy consumption over time and specifies the place of natural gas in it. It also presents the current state of energy consumption in the world. Then, we will expose the available econometric tools for demand modeling and the difficulties associated with this particular exercise, and finally, we will present the contribution of this thesis in the field of energy demand forecasting.

Energy consumption in History

According to Malanima (2014), human history can be separated into two major eras: the 5-10 million years between the birth of the human species and the beginning of the modern era, i.e. about 5 centuries ago, and the recent history of the last 500 years, which has seen a rapid acceleration of energy consumption.

Pre-modern Organic Vegetable Economies

At the end of the 18th century, there were three main economic sources of energy. The first source was food, the second was firewood and the third was fodder for working animals. A relatively small contribution came from two other vectors: falling water, whose potential energy is exploited by water mills, and wind, both used sailing ships. We can split this pre-modern period into three epochs: food, fire, and agriculture. For the first era: Food, which goes back to the beginning of humanity, more precisely to the first evidence of the use of fire by humans. This major discovery refers to several different regions of the world and can be dated from one million to 500,000 years ago. Human energy comes only from food, which corresponds to about 1,500 to 2,000 calories per day (Malanima, 2012). The discovery of fire pushes humanity to a new level of energy consumption. It is important to point out that coal, oil, and natural gas, the basic fossil-fueled sources of today energy is based on the same principle (oxidation) for heating or light. Fossil-fueled sources are carbon compounds like bread or firewood. Although with fire, the calories per head increased drastically from 2,000 to 3-4,000 per day or more, the efficiency of its use was very low. A fire in the open air cannot provide man with more than 5 % of his calories, i.e. no more than 2 to 300. The useful energy exploited by the population from food and fire did not exceed 500-700 Calories.

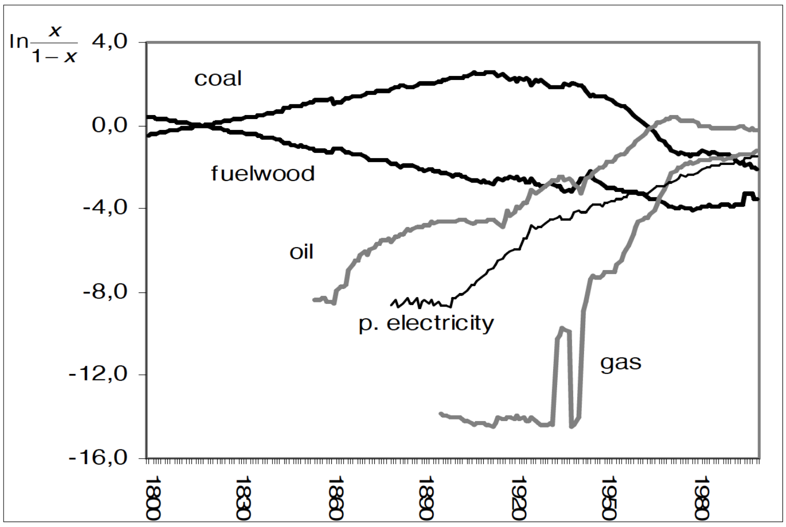
The third era was agriculture-based this was a kind of revolution. Several innovations allowed a more efficient utilization of man's power, during the first period (from about 5000 years until 3000 BC). The fundamental change was represented by the taming of animals (oxen, donkeys, horses, and camels), and their use in agriculture and transport. During this period, fuels are more animals-based; e.g. the wheel, the working of metals, pottery, the plow, and the sail increase. The agrarian civilizations' main innovations in the realm of energy were water and windmills which was invented respectively 3 centuries BC and in the 7th century AD. Despite several important differences exist among the three ages, there is also one major common denominator, which is the high-level dependency of man to the environment. The dependence of this energy system on soil implies several constraints to the possibilities of economic development. This pre-modern organic vegetable economics have some similarity, they are based on renewable sources. Organic vegetable economies have been sustainable since solar energy has allowed a continuous flow of exploitable biomass. The availability of more plant-based sources involved the extension of arable and grazing land and the collection of firewood, which is difficult to transport over long distances. The production of organic vegetable sources and economics of the pre-modern Organic Vegetable Economics is also highly dependent on climate.

Modern Organic Fossil Economies

At the start of modern growth around 1800, water and wind were the only non-organic sources. The study of this first energy transition which took place in the early 19th century in the UK is very insightful about the drivers of our current and unachieved energy and economic transition from fossil fuel economics to Renewable Energy Sources (RES) based economics. This first energy transition was based on fossil sources, coal, oil, natural gas, which were also products of photosynthetic processes, such as food and firewood. Their formation took place during the Carboniferous era, about 300-350 million years age (See Smil (2015) for more details about the formation of natural gas formation in the underground). Coal was widely extracted, by English workers for domestic uses. There is an academic consensus about the origin of this first energy transition which was based on technical innovation. It was during the 18th century, with the invention of the steam engine by Thomas Newcomen and James Watt. Indeed the fundamental technological obstacle that had for millennia limited the capacity of the economic systems to perform work, was only then overcome. This innovation was one of the main drivers of the first industrial revolution. The reason for the spread of this technology is more subject to debate. According to Malanima (2014), the spreading of James Watt's machine is due to the scarcity of the main converter of the organic vegetable energy system, land. Malm (2013) attempts to address the dynamics of the fossil economy by examining the causes of the shift from water to steam in the British cotton industry during the second quarter of the 19th century, which is based on the diffusion of the steam engine. Contrary to Malanima (2014), it shows that the shift as driven by scarcity is refuted, "the choice of steam was motivated by a rather different concern: power over labor." Both Malm (2013) and Malanima (2014), agree about the fact that "Although not sufficient condition of modern growth, energy transition was a necessary condition", they rely on different approach is to explain the link between this transition, and the modern fossil fuel growth. Malm (2013) open a dialogue with Marx on matters of carbon and outlines a theory of fossil capital. The link between energy consumption and modern growth, is subject to debate (for a more historical perspective of this debate see Malanima (2012)). The Watt's machine was the first, but technical progress has created others, such as Parson's, which had different applications and have other economic impacts, for example, the development of maritime trade (See Smil (2006)). After this first energy transition, energy consumption per capita decreased in Europe during the 18th century, while from 1800 to 2005 it increased much more than the population (See Smil (1994)). Until roughly 1840, energy consumption per capita did not increase in Europe, because the supply of fossil fuels grew at the same rate than the population. From 1840 onwards, growth was rather remarkable until the First World War. After a period of stability between the two World Wars, a significant increase took place from the 1950s to the 1970s, followed by a slower increase. In the long term, growth in per capita energy consumption is almost constant, with brief deviations due to wars or periods of rapid economic growth. We now describe the energy substitution paths that lead to the place of natural gas consumption in our economy, Since this initial discovery of James Watt's machine.

From the end of 19th to the beginning of the 21st century

Figure 1: Shares of any fuel on the total fuel consumption in Europe 1800-2000 (ln)



Notes: x, on the vertical axis, refers to the share (%) of an energy carrier on the total of the 5 energy carriers. *Source: Malanima (2014).*

Coal

As the world's use of coal eventually outstripped that of wood and crop residues, the balance likely tipped in the 1890s. England entered the 19th century as a coal-dominated economy (in 1800 it produced 80% of the world's coal). Most of Western and Central Europe made the transition before 1870, while the United States derived more energy from coal than from wood until the early 1880s, and in Japan and Russia wood dominated until the end of World War I. Coal fueled the transition from traditional artisanal economies to modern mass manufacturing and made steam engines the most essential engines of the industrial revolution. Coal's share of overall commercial energy use was 95% in 1900, and did not fall below 50% until the early 1960s. But as the relative importance of coal declined, its absolute production increased. From the 2000s onwards, world coal production again exceeded four billion tons. Coal provided just over 23% of the world's primary energy needs. In the early 1960s, crude oil became the world's most important fossil fuel, and natural gas consumption grew faster than oil consumption.

Electricity

The combustion of fossil fuels and the use of high-temperature (pressurized) steam to operate turbine generators have been the most common means of generating electricity since the commercial beginnings of the industry. In 1882, Thomas A. Edison built the world's first two small coal-fired power plants in London and New York City. The rapid growth in capacity of typical power plants was made possible by a combination of several key inventions: steam turbines, transformers, the conversion of direct current to alternating current, and high-voltage transmission, as well as a process of continuous innovation and efficiency improvement. The combustion of fossil fuels now produces about 63% of the world's electricity, and the best efficiencies in the whole process are about 40%. The other three main sources of electricity are hydro, nuclear and renewable energies for the remaining 37%.

Oil & Natural Gas

Like coal, crude oil was discovered early in Antiquity. Here we will see a second energy transition taking place rapidly from the 1950s onwards. Coal remained the dominant fuel in post-World War II Europe throughout the 1950s, but the continent's subsequent transition to an oil and gas-based economy was rapid (See Figure 1). This transition is not based, unlike the first one, on technological innovation, but rather on the double efficiency of coal substitutes. Firstly, the energy efficiency of gas and oil is much higher than that of coal. Second, the substitutes are easier to transport. This recent transition has a direct impact on the theory on non-renewable resource associated with oil and gas, which is still quite close to a resource theory (See Jevons

(1865)). From a geographical point of view, before World War II, the United States were the only industrialized country with a significant share of hydrocarbons in its primary energy supply. As Western Europe did not have significant sources of hydrocarbons, it became a major importer of crude oil from North Africa and the Middle East. The discovery of the giant Groningen natural gas field in the Netherlands and the development of oil and gas fields in the North Sea (the first productive wells were drilled in the 1960s) and imports of Siberian gas into Central and Western Europe (in the early 1980s) accelerated the continent's transition to hydrocarbons, leading to either the complete disappearance (in the Netherlands) or the drastic reduction (in France, the United Kingdom, and Germany) of coal mining.

For natural gas, many technical improvements appeared during the post World War II period in the following fields: drilling (offshore techniques) and transports (pipelines). Worldwide exploration for hydrocarbons has dramatically changed the distribution of oil and gas reserves. The first Middle Eastern super-giant oil field was discovered in 1938. The world's largest oilfield (containing almost seven percent of the world's oil reserves), was drilled in 1948. World civilization will depend on crude oil for decades to come while continuing to use more natural gas, the simplest, cleanest and in many ways the most desirable of all hydrocarbons.

We will now focus, on the history of natural gas which contrary to the mainstream thought is quite old. Indeed the only well-documented use of natural gas in pre-industrial society was its combustion to evaporate brine in the landlocked Sichuan region of China, which began in the early Han Dynasty (about 200 B.C.). Natural gas is composed mainly of methane, with small amounts of ethane, propane, hydrogen sulfide, and nitrogen. It is often blended with crude oil in the same tank. During the first decades of the oil industry (when there were no long-distance, high-pressure pipelines), this so-called associated gas had to be burned, if it could not be used locally.

Gas turbines have also found a variety of stationary applications: they are the preferred choice for powering the large centrifugal compressors that push natural gas through pipelines and provide the pressure needed for many chemical and metallurgical processes, and they are increasingly being used to generate electricity in relatively small, decentralized facilities.

Technical improvements have raised the efficiency of these machines to over 40% and, in combined cycles (using the hot gas leaving the pipeline to heat water in a smaller steam turbine), they were the first converters to achieve efficiencies in excess of 60%. At this point in this introduction, we will explain the place of natural gas consumption throughout history until the beginning of the 21st century, in the following paragraphs we will detail the geographical specificities and characteristics of gas consumption in the contemporary world.

Contemporary Natural Gas Demand

Figure 2: Global energy mix (%)

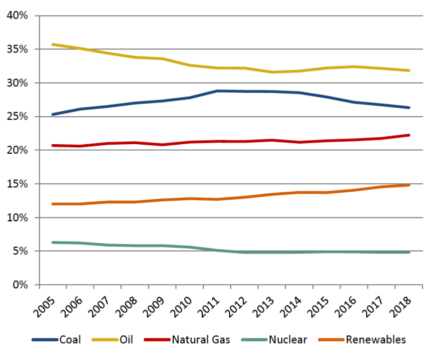
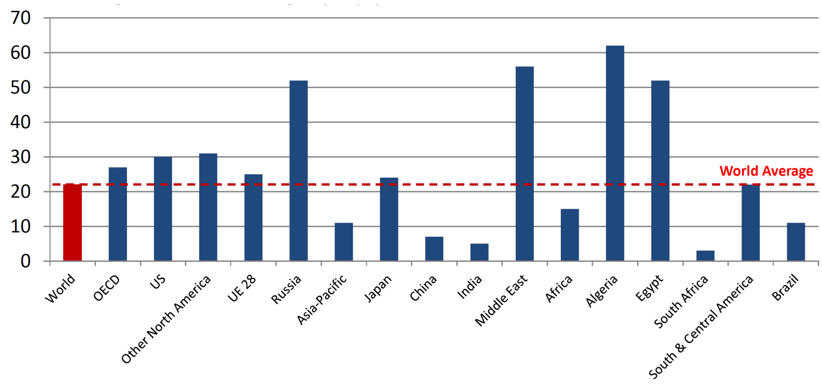


Figure 3: Global energy mix (%) by world zone



Notes: *Data for 2018. Source: CEDIGAZ (2019)*

Figure 2, shows that, as a percentage of the global energy mix, natural gas is ranked third, behind, coal, and oil. It is interesting to note that natural gas is the only fossil fuel whose share in the energy mix is growing. It is also ahead of renewable energies in terms of share in the global energy mix. Figure 2 is also interesting because it highlights for the first time the duality of natural gas, its strong link with oil, and its complementarity with renewable energies. Nevertheless, Figure 3 shows us the disparities in the share of natural gas in the energy mix of the different countries. What has been noticed is that in the world areas with a high coal energy mix (Asia-Pacific, China, India, and South Africa) the share of natural gas is below the world average. Also, the crude oil-exporting countries (Russia, US, Middle East, Algeria, and Egypt), have a large share of their mix based on natural gas. The reason is twofold, natural gas is initially considered as a co-product of oil, and the second is that gas transport requires large investments in terms of infrastructure.

Figure 4: World map of natural gas consumption in 2018

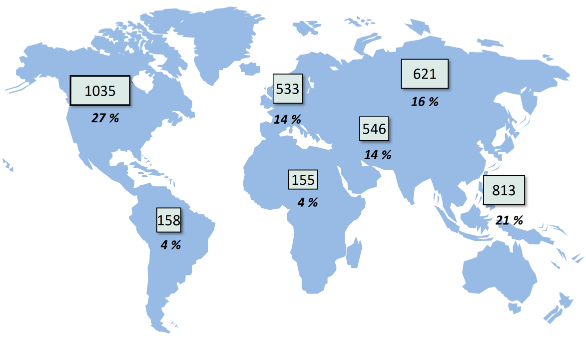
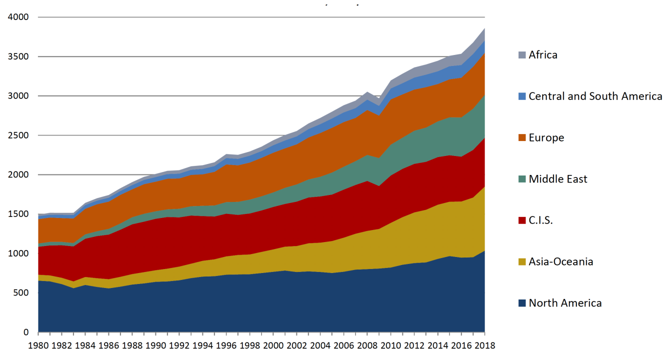


Figure 5: Evolution of natural gas consumption in world



Notes: *Data for 2018. Source:* *CEDIGAZ (2019)*

World gas consumption is mainly concentrated in three major markets: North America, Europe and Asia, which together account for 78% of total gas consumption, which will reach 3,681 billion cubic meters in 2018.

The European market

The European market has a consumption of 620 Gm^3 in 2018. It is the third-largest consumer region. The main driver of gas consumption in Europe is the electricity sector. Europe is also the first exporting region (50% of inter-regional exports). Europe has a long gas tradition and can be considered as a mature market: gas provides 54% of primary energy needs, ahead of oil (20%), and coal (17%). Russia is the second-largest consumer market in the world (450 Gm^3 in 2018) as well as the first export market (249 Gm^3). There is current competition from American gas shales mainly on LNG (the acronym LNG, for Liquefied Natural Gas) exports (but Russia maintains its monopoly on pipeline exports). As the evolution of gas demand in 2016 is identical to the level of the early 2000s (effect of the economic context and energy savings), there has been a recovery in consumption since 2016.

The Asian market

In contrast to the EU, the Asian market can be regarded as a relatively young market with strong growth potential (emerging markets). Nevertheless, there are many disparities between national markets (market maturity, the share of gas in the energy mix, network density, regulatory context). In Asia, two national markets dominate regional consumption: China (33%) and

Japan (15%). Due to the strongest pressure from the international community to comply with CO_2 emission constraints, there is a general trend in Asia towards a growing role of natural gas in the power sector (China) and industry, petrochemicals and fertilizer production (India, China, South-East Asia). There is an important development of gas in the residential-tertiary and transport sector (CNG/LNG), linked to urbanization (China, India). Asia can be seen as the engine of growth in global gas and LNG demand.

The North American market

The North American market has a consumption of 1039 Gm^3 in 2018. It is the region that consumes the most. Since the shale gas revolution, there has been a shift from being an importer to an exporter in 2018. The market can be considered an integrated market. There are a long gas tradition and a mature market: gas provides 31% of primary energy needs, after oil (37%) but ahead of coal (14%). Sectoral consumption is dominated by the electricity (35% of gas consumption) and residential/tertiary sectors (25%). The problem of the continental state (the USA or Canada) is that it requires long distances between places of production and consumption for supply, which has a high cost in terms of infrastructure. There is a very dense and interconnected transport network. For linking supply and demand, the North American market has opted for a deregulated and liberalized market (Henry Hub).

Natural gas consumption is not uniformly distributed (see Figure 4), it is more concentrated in OECD countries, and/or in oil-exporting and producing countries. We will set aside Russia, which has a somewhat special status. It is interesting to separate the two families, the use of natural gas in the energy mix is not the same. It's depending on the state of completion of the energy transition and the endowments of soil resources. In the case of countries that have "completed or well advanced" in the energy transition, gas is positioned with a so-called peak demand, i.e. gas-fired thermal power plants are only activated when there is a high demand for electricity consumption. The case of France is characteristic of this type of country which has succeeded (thanks to its ultra-nuclearization) in "completing" its energy transition (low CO_2 emissions). The first chapter (See Chapter 1) of this thesis talks about the modeling of natural gas demand in this case. Figures 4 and 5, show that the United States is the leading consumer, since the Shale gas revolution it has also become the world's leading exporter of natural gas (in its LNG form). The demand for natural gas in the US is still growing (See Figure 4). This leadership position, coupled with a liberalization of the markets in the United States, pushed us to study the United States in this thesis, two chapters (See Chapters 2, 3 and 4) of this thesis are devoted to the US.

Contemporary Natural Gas Inventories and Transport

Because of Natural gas requires a volume 1000 times greater than oil for the same energy content, this requires the development of local and then regional markets (North America, Europe, Asia). Transportation is mainly carried out in two ways: by pipeline, a pipe used to transport large quantities of fluids (oil, natural gas, etc.) over long distances, and by LNG refers to natural gas that is kept in liquid form. At atmospheric pressure, when natural gas is cooled to a temperature of about -161°C, it turns into a colourless, odourless, non-corrosive and non-toxic liquid. In both cases, natural gas transportation is very capital-intensive. As a comparison, the investment cost for a gas pipeline is 2 to 6 G$/1000 km, while the investment cost for an LNG liquefaction system is 500 to 1500 $/t. In fact, the two transport systems are used differently and their costs vary according to the distance (See Figure 6). LNG is profitable the farther away you go. It is important to specify that with both systems it is technically possible to expect a globally integrated market. Figure 7 shows us that pipeline transport is still in the majority (77% in 2018) but that the share of LNG is still evolving. It is important to specify that this thesis does not focus on LNG.

To properly model natural gas the demand, it is essential to have a good knowledge and good modeling of what is called residual demand. This is the difference between what has been produced and what has been consumed. This residual demand can be found mainly in two places: in the variation of inventories and the flexibility linked to the transport of gas by pipeline. It could be also in LNG transport vessels.

Figure 9, shows the evolution of this residual demand over time. That is what happens

Figure 6: Evolution of Natural gas transport cost

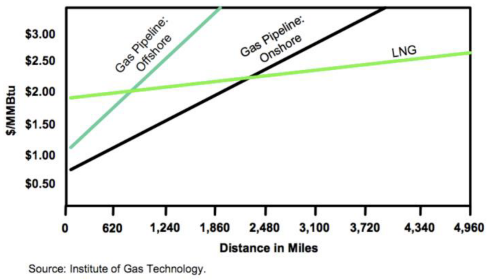
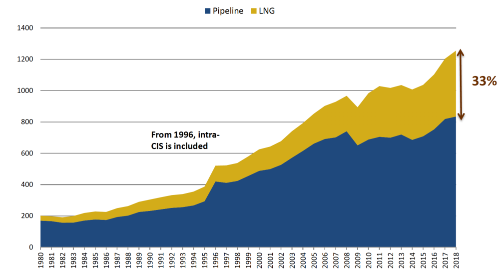


Figure 7: Part of pipeline and GNL transport in the world.



Notes: *Data for 2018. Source:* *CEDIGAZ (2019)*

Figure 8: World natural gas inventories

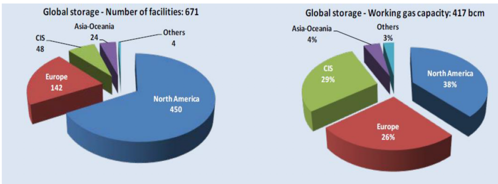
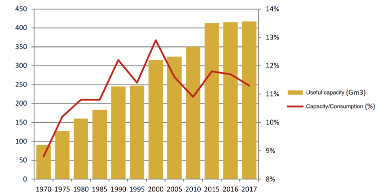


Figure 9: Evolution of Natural Gas storage capacity



Notes: *Data for 2018. Source:* *CEDIGAZ (2019)*

when the red line is above the yellow histograms, we draw on the stocks, in the opposite case we refill stocks. Gas storage is mainly used to adjust supply to demand. Gas storage allows modulation at different frequencies: Year - Season - Week - Day - Hour. This flexibility is a real asset in favor of natural gas in the fight against climate change, as it allows us to manage the electricity peak demand. This flexible storage mitigates the strong variations in demand linked to the growing role of gas as a supplemental source to renewable energies (intermittent electricity generation). This case is studied in the chapter 1 of this thesis with an application in France.

Natural storage also helps to ensure the security of supply (strategic storage). It allows an optimization in operating the pipelines to increase efficiency (maintenance of a high load factor) which leads to a reduction of the supply cost. Figure 8, show that the most important area in the world in terms of gas storage capacity is North America (US+Canada). This is why in this thesis we are also interested in modeling the variations of the residual demand for natural gas in the US (See Chapter 2). The management of the storage capacity that can be contained in the linepack[1], it requires a great capacity to forecast gas consumption. Indeed, this residual demand contained in the stocks has flexibility but requires significant time inertia. This is why the chapter 1 of this thesis is also interested in forecasting the aggregate demand for natural gas in a country that has relatively little available storage capacity, France.

A brief introduction to supply and hedging strategies

In 1961, international trade in natural gas was limited to imports from the United States, to Mexico, Canada, and some deliveries from the former USSR to Poland, i.e. barely 1% of world production. It was not before 1964 and the first commercial delivery of Algerian LNG to the United Kingdom that another inter-regional route was created. In 1972, the first cubic meters of gas crossed the borders of Western Europe. Since then, gas trade between countries and continents has continued to grow, accompanied by the development of new, sophisticated, and expensive transport chains. Growing regional imbalances between consumption and production areas explain the very strong growth in international trade in natural gas from the 1970s onwards. LNG has seen its importance in trade increase significantly, to transport natural gas over long distances, and to promote globalization and trade flexibility.

The North American market is considered as a liberalized market. In Europe, the objectives of market liberalization are: to ensure the free movement of gas and to enhance the security of supply and industrial competitiveness of European energy companies through the free functioning of the market and to create a single European gas market, integrated and competitive, flexible and non-discriminatory, with market-based supply prices. In the rest of the world, they exist more local and monopolistic market which are interconnected by LNG. The historical, pre-liberalization principle of supply is as follows: a firm (or country) producing natural gas contracts on a long-term basis with the local firm (trader) responsible for ensuring supply in its country of consumption. Long-term contracts work in the following way: the producer and the trader contract on a volume, a sale/purchase price, and a validity period (generally from one to ten years). Most of the time, the feature of the contract is called the Take-Or-Pay (TOP). The volume must be exchanged since if the trader does not wish to acquire the natural

[1] Refers to the volume of gas that can be "stored" in a gas pipeline.

gas, he will still have to pay for it. The long-term contract ensures the trader a secure supply, which will enable him to meet consumer demand. Long-Term Contracts create a sharing of risk between the importer (volume risk) and the producer (price risk).

There are other types of contracts. These different types of contracts are detailed in Von Hirschhausen and Neumann (2008). The evolution of the spot price of natural gas seems to be fixed by the price of oil. Indeed these two commodity prices are highly correlated. This correlation can be explained by two phenomena: the establishment of long-term contracts, on the supply side, and energy substitution, on the demand side, particularly for electricity production. Currently, the price of long-term contracts between producers and traders is set based on a formula indexed to the price of oil.

Von Hirschhausen and Neumann (2008) shows that the duration of long-term contracts is declining for recently renewed contracts. The liberalization of natural gas markets introduces new tools to manage the risk for both traders and suppliers. This long term contract is less and less exchanged Over-The-Counter but in new future exchanges market place (such as Henry Hub). These contracts called "*futures*" have a quotation, price, and delivery date. Natural gas market participants can use new hedging strategies tools to cover the risk. Hedging strategies involve buying or selling futures to deal with the risk in the price of the concerned commodity. The commodity-specific stylized characteristic of each underlying commodity - the convenience yield - was first introduced by Kaldor (1939). It represents the flow of services for the physical holder of the commodity but not for the owner of a written futures contract on the commodity. Commodity convenience yield relates to current and expected future market conditions and thus determines storage decisions.

The future of energy demand in energy transition: the key role of natural gas

Preventing climate change is one of the priorities of world energy policy, which aims to massively reduce its greenhouse gas emissions. By 2030, it aims to reduce greenhouse gas emissions by at least 40% compared to 1990 (in the E.U). In this context, the future place of natural gas in the energy mix remains uncertain. If methane shares with oil or coal the lack of being a fossil-fuel, exhaustible, and greenhouse gas generating energy source, some arguments suggest that this energy source can also make a useful contribution to the success of the energy transition. The substitution of oil and coal by natural gas reduces CO_2 emissions. Helm (2012) shows the energy transition needs to be achieved by shifting from coal to natural gas for electricity generation. This storable and flexible energy source is complementary to intermittent renewable energies in electricity production. Paltsev et al. (2011) shows that the outlook for gas over the next several decades is very favorable. In electric generation, given the unproven and relatively high cost of other low-carbon generation alternatives, gas is likely the preferred alternative to coal.

Paltsev et al. (2011) focuses on uncertainties including the scale and cost of gas resources, the costs of competing technologies, the pattern of greenhouse gas mitigation, and the evolution of global natural gas markets. Natural gas will be a bridge fuel. More precisely natural gas can be an effective bridge to a lower emissions future, but investment in the development of still lower CO_2 technologies remains an important priority (Paltsev et al., 2011).

Of course, natural gas remains fossil-fueled energy, and it is ultimately necessary to get out of it. This is gas phased out: Zenghelis et al. (2014) shows a transition from coal to clean energy is to be made via intermediates (for example, gas), the use of gas (without carbon capture) should be agreed upon to be on a time-limited basis. Further, to avoid gas lock-in, research in fully clean technologies would need to be strongly stepped up over the intervening period, along with other supportive policies. Holz et al. (2016), analyzes the infrastructure needs of the European natural in response to the decarbonization of the European energy system. They investigate pathways of future natural gas consumption: i) a decreasing natural consumption, ii) moderate increase of natural gas consumption, along the lines of the IEA's New Policies Scenario; and iii) a temporary increase of natural gas use as a technology, followed by a strong decrease after 2030. Holz et al. (2016) show that rent import infrastructure and intra-European transit capacity are sufficient to accommodate future import needs in all scenarios. The existence of this mature gas infrastructures makes it possible to accelerate the deployment of technologies to produce methane from renewable sources (e.g. biogas or Power-to-Gas technologies to exploit surplus electricity from intermittent ENR sources). Faced with these challenges, it is striking to note that our knowledge of both gas and oil demand modeling remains limited because it is largely based on old empirical work and methodologies that are now dated. **Therefore, the scientific objective of this thesis is twofold. First, we analyze quantitatively the economic determinants of energy demand. Second, we develop new forecasting models. The aim will be to develop and implement modern investigation methods to study the fundamentals of energy demand. This will involve measuring the reactions of this demand to changes in different variables such as market participants behaviour, energy prices, and also real-time variables such the temperature, and assessing their dynamics. On the basis of this finding, we develop forecasting models for both demand and supply. This theme is currently undergoing a very important renewal due to the concomitance of three phenomena: (i) the provision of new, enriched and detailed databases that allow for a finer analysis of the determinants of energy demand; (ii) increased societal demand as a result of debates on energy transition policies and the need to calibrate the instruments associated with these public policies; and last but not least, (iii) the development of more sophisticated analytical methods, to manage infrastructures and consumers behavior in the context of the energy transition.**

The econometrics of energy demand

There is no "technique for all seasons", it is a matter of choosing the methodology whose strengths best suit the task at hand. The scope of any analysis is fundamentally limited by the availability and quality of the data. Hartman (1979) lists three main types of decisions faced by energy consumers: the purchase (or replacement) of energy using a durable good providing certain services (heating, lighting, industrial transformation); the choice between equipment of different technical and energy characteristics providing the desired service; the purchase and installation of the equipment, the choice of the intensity of use (capital use). These decisions cover the short term when stocks of energy-using goods are fixed, and the long term, when the stock can be changed. They have implications for the specification of an appropriate model.

The first classification of an energy demand model can be found in Labys (1982): (1) econometric market models; (2) econometric process models; (3) econometric demand models; (4) spatial equilibrium models; (5) optimization models; (6) system dynamics models; (7) cost-efficiency based models.

Econometric market models are based on the assumption of a market-clearing price principle. Most of energy demand models are not based on this principle, except when energy markets are linked to a macro-model. The models are mostly static and are based on the seminal work of Hudson and Jorgenson (1974) and Fraumeni and Jorgenson (1981). In this type of models, the economic variables and prices are endogenized. This kind of model has a full description consistent with economic theory. The problem is data calibration and dynamic adjustment. These models are often based on modeling the process from the resource to the final product. Demand is often exogenous and prices are not compensable by the market. (Case of the TIMES/MARKAL or LEAP models). The demand models in the Labys (1982) classification are limited to the demand side, not the supply side, and not to the production theory.

Spatial equilibrium models focus on the inter-regional efficiency of production, distribution. The optimization model has been used to optimize the econometric models for forecasting purposes.

The system dynamics models are based on the attributes of commodities that lend themselves to systems engineering or analysis. Cost-efficiency based models are essentially accounting models and may involve the evaluation of energy costs.

This thesis focuses on market, process, and demand econometrics. The weaknesses of these models is, to include specific engineering decisions, provide long-term forecasting, abilities to the analysis of structural changes, include non-exogenous government impact, model projection of technological evolution, and irreversibility. For a full discussion of the advantages and disadvantages of each model, see Table 2 in Labys (1982). This thesis provides insights into market-clearing prices in energy demand modeling. These prices result from the interaction of the supply and demand functions. Given a time series of these prices, the classical identification problem is to know which component - demand or supply - would be approximated by the data if each function were estimated. The correct approach is to estimate the demand and supply functions simultaneously. For a general discussion on this models, see Bohi (1981); for theoretical aspects, see Wonnacott and Wonnacott (1979). The chapter 1 is in the same vein of this literature.

Modelisation of energy demand could be estimated at different levels: aggregated level (the case of this thesis) or multi-sectoral levels (for example, transport, industry). This difference of scales in modeling generates constraints to be modeled. Some consumption sectors are characterized by strong energy substitution. The latter occurs when consumers have the possibility of changing, modulo certain inertia of use, the energy source they consume, according to the market prices of the different fuels to which they have access. The industrial sector, for example, shows strong energy substitution. If such substitutions exist, they are likely to create competition between energies, within consumption itself. For example, expensive energy would thus become relatively less used. On the contrary, inexpensive energy would see its consumption to increase over time. Thus, in the equilibrium, the prices of competing energies are correlated. From a purely economic viewpoint, this explanation invokes the short- and long-term cross-elasticities (hence the need to calculate them) between the prices of different energies. Some end-use sectors, such as transport, show very little substitution. This sector currently accounts for around 20% (WEO, 2018) of total consumption in France, which leaves 80% of consumption likely to be subject to more or less strong substitution. Therefore, the estimation of the cost of each energy is complex. It must take into account the market price, the relative costs of capital investment, and the cost of CO_2, which makes it possible to take into account at least partially the impact of environmental policies.

Forecasting demand model

During the years 1975-1995, much work has been devoted to empirical modeling of energy demand. Forgotten for a while, this theme is currently undergoing a very important renewal (Ryan and Plourde, 2009; Medlock, 2009). Suganthi and Samuel (2012), review all types of energy demand forecasting models, for different types of energy (electricity, coal, oil, and natural gas), geographical areas, and granularity of demand. It should be noted that some of the econometric techniques derived from the economics of other energies might easily be transposed to the modeling of natural gas demand, but for the sake of clarity we have reported only the one concerning natural gas (for other applications see Suganthi and Samuel (2012)).

Suganthi and Samuel (2012) defined econometric models such as time series models are simple models that extrapolate the future energy demand.

Among these models, we can discern the basic univariate models such as: Hunt et al. (2003) investigated the energy demand on a sectoral basis for the UK. Kumar and Jain (2010), have employed a Markov model to forecast crude oil consumption with a rolling mechanism to forecast coal, electricity consumption and spectrum analysis to predict natural gas consumption. Aras and Aras (2004) simply used a first-order autoregressive time-series model to predict the natural gas consumption. Regression models have been used to forecast the coal, oil, gas, electricity demand see Farahbakhsh et al. (1998); Sharma et al. (2002). Main econometric models try to correlate the energy demand with other macroeconomic variables, such as: McAvinchey and Yannopoulos (2003) implemented a three equation model for energy modeling and forecasting energy demand in the UK and Germany. An econometric model is then used to correlate the price of electricity, oil, gas, coal, total energy demand, and technological progress. This multi-system equation model is corrected with an error term for the long-term equation. A modified logit function model is used for extrapolating crude oil and natural gas demands for France Mackay and Probert (1995), in this model, population and GDP/capita are useful in forecasting the demand. Econometric models are also developed for the various petroleum products and natural gas for India Parikh et al. (2007). Variables such as GDP/capita, population, price are considered to forecast the demand. Eltony (1996) uses an econometric model to forecast and study the determinants of the natural gas demand in Kuwait. Ang (1995) reviewed a list of studies related to industrial energy decomposition. Two common approaches for decomposition are the energy consumption (EC) and the energy intensity (EI) approaches. Cointegration models have to be considered, Smith et al. (1995) provide an econometric model of fossil fuel demand for eight OECD countries, relating coal, oil and gas demands to GDP and prices. Also, a model that includes endogenous technical progress has been estimated, aiming to include both innovations in energy and structural change in the economy as long-term determinants for energy consumption. Lee and Chang (2005) examines the stability between energy consumption and GDP for Taiwan. Aggregate and disaggregate data of energy consumption, including coal, oil, gas, and electricity, is used. Unit root tests and the cointegration tests allowing for structural breaks are performed on the data. ARIMA models have been extensively used in energy demand forecasting. (Ediger et al., 2006; Ediger and Akar, 2007) develop a decision support system for forecasting fossil fuel production using regression, ARIMA, and SARIMA method for Turkey. Erdogdu (2009) has estimated short and long-run price and income elasticities of sectoral natural gas demand in Turkey.

As mentioned in the introduction, the number of readily available data increases significantly. More and more model includes, real-time data to most closely approximates the demand. For example, Elkhafif (1996) presents an iterative econometric technique for energy forecast which corrects the non-normal weather conditions data. The model is applied for sectoral natural gas data for the province of Ontario, Canada.

This new massive input of data, leads to the birth of more and more complex models to extract (Machine learning) the key determinants of our variable of interest and to try to predict it. That's why expert systems and neural networks were being used extensively for energy demand forecasting. Gorucu (2004) uses Neural Networks to predict oil and natural gas consumption. Gorucu (2004) have used ANN to predict the gas demand for Ankara. Nguyen and Nabney (2010) also use wavelets techniques for short term load forecasting. An input-output model is used to assess how social and economic changes will affect energy requirements and energy intensity. Canyurt and Ozturk (2008) develop a genetic algorithm demand estimation models (GA-DEM) to determine the future requirement of coal, oil and natural gas in Turkey based on population, gross national product, import, and export. Two methods are used to estimate the logistic parameters- one using nonlinear programming (NLP) and the second using a genetic algorithm (GA). Integrated models-Bayesian vector autoregression, Support vector regression, Particle swarm optimization models are also widely used. Particularly Crompton and Wu (2005) uses the bayesian vector autoregressive model to predict energy demand for China more precisely the primary energy requirement of coal, oil, gas, hydro is projected till 2010.

For oil, reference models are the structural model based on the VAR methodology is developed in (Kilian, 2008a,b, 2009; Kilian and Murphy, 2012; Baumeister and Peersman, 2013) with a recent debate: How to measure economics activity Kilian (2019); Baumeister and Hamilton (2019a); Baumeister et al. (2020). The associated forecasting performance is studied in (Baumeister and Kilian, 2012, 2014; Baumeister et al., 2014; Baumeister and Kilian, 2015; Baumeister et al., 2017b, 2020). Our work is in the same vein of these structural model. The predictive role of future prices and the link with inventories is studied in Alquist and Kilian (2010a); Alquist et al. (2014); Fattouh et al. (2012); Kilian and Murphy (2014a); Kilian and Lee (2014).

In this renewed context, our research focuses on energy demand determinants. The broad goal of this dissertation is to provide newly considered advanced econometric tools and approaches to forecasting energy demand in EU countries and the

US with a focus on natural gas.

The purpose of the first chapter: **"How are day-ahead prices informative for predicting the next day's consumption of natural gas? Evidence from France"**[2] is to investigate, for the first time, whether the next day's consumption of natural gas can be accurately forecasted 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 and apply it to the case of France. Our results document the existence of a long-run relation between demand and spot prices and provide estimates of the marginal impacts that these price variables have on observed demand levels. We also provide evidence of the pivotal role of the spark ratio in the short run 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.

The second chapter: **"Considering real-time demand to forecast the U.S. natural gas price in real-time: The role of temperature data"**[3] provides evidence of the pivotal role of temperature data to forecast natural gas prices at the Henry Hub in real-time. Considering a newly constructed temperature index as an additional exogenous variable in a Bayesian vector autoregressive (VAR) framework significantly increases the forecast accuracy at horizons up to 12 months. Our approach is new to the energy price forecasting literature as it considers both supply and demand at the same time and further includes the temperature as a proxy of real-time demand of natural gas with superior forecasting results.

The third chapter: **"A Structural Non-causal VAR Model of the Global Oil Market: the Role of Oil Supply News Shocks"**:[4] Nonfundamentalness issue on the global oil market

[2] This chapter is based on Thomas et al. (2019) accepted for publication in *The Energy Journal*
[3] This chapter is based on Moussa et al. (2020)
[4] This chapter is based on Moussa and Thomas (2020)

has been addressed by either augmenting small-scale VAR models by additional variables or latent factors, or using external instrument or proxies leading to more credible identification scheme. We tackle this problem by estimating a non-causal VAR model for standard global oil market variables. We identify the oil supply news shock as a shock that drives global oil production the most for a finite time horizon. First, Our findings highlight the prominent role of expectations in propagating the shock. Second, we show that a negative oil supply news shock results in abrupt and permanent reaction in global oil production, global economic activity and in oil inventory. However, the oil supply shock has only a limited effect on oil price. Finally, news shock about oil supply shortfalls do have macroeconomic consequences as it causes a substantial decline in US industrial production.

The aim of the final chapter: **"The role of expectations in predicting the real prices of oil: a non-causal analysis"**[5] revisits the predictive power of convenience yield for oil by incorporating expectations into an empirical specification through the estimation of Bayesian non-causal VAR. We empirically show that expectations play a significant role in the determination of oil prices. Second, we provide empirical evidence that real-time forecasts of real oil prices can be remarkably more accurate than the no-change forecast and significantly more accurate than real-time forecasts generated by existing structural models relying on Bayesian VAR.Beyond the traditional analysis at the monthly frequency, we further investigate the forecasting accuracy of our empirical specification at the daily and weekly frequency, resulting in interesting findings for potential investment purpose.

Finally, in the conclusion we summarize the main findings of this thesis and puts forward several objectives for future research.

[5] This chapter is based on Thomas (2020)

Chapter 1

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

[This chapter is based on Thomas et al. (2019) accepted for publication in *The Energy Journal*]

1.1 Introduction

The accuracy of gas demand forecasts issued by Transmission System Operators (TSOs) is now becoming an important matter 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,

[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 (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).

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 using the information in 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 to significant differences in the perceived degree of trading liquidity observed at the European gas hubs (Heather and Petrovich, 2017). Thus, a fundamental public policy issue is 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 examining the informational content of day-ahead electricity prices. They hypothesized that if day-ahead markets for electricity were 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 only the information contained in the day-ahead price series. The authors examined

[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.

California's PG&E aggregation area and applied 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 major informational content of day-ahead prices in the case of electricity.

Our paper is 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 (2014)'s approach in highlighting what necessary specific dimensions must be considered to produce accurate natural gas demand forecasts. However, we acknowledge that dealing with gas prices requires modeling specificities that differ from the ones 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 time series (unit-root) properties of the data. We further elaborate on these three distinctive features in section 1.2 as they provide the essential justifications for our modeling choice and are thus key in our analysis. 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 the Threshold ARDL (TARDL), that we propose here, and that is a straightforward extension of the genuine ARDL model. 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]

Our application deals with 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 information about the weather 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 period 2015-2018. 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 reaction of daily gas consumption to the price of natural gas in the short run. 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 producing day-ahead forecasts that are considerably more accurate than those published by network operators.

Our main contribution to the literature is to show that publicly available information, such as day-ahead prices, can be used to produce efficient forecasts of tomorrow's consumption by relying on a quite simple 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 literature in 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 is much more relevant for eliciting the 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 some kind 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 marginal impacts that gas and electricity prices have on natural gas demand in both the long- and the short-run, and thus on the dynamics of the reactions of natural gas demand to energy price changes. As these marginal impacts play an important role in the models developed to examine the effects 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 consumption 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

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

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, see footnote 1 for more details).

The remaining sections of this paper are organized as follows. In the next section, we present some attributes that are typical for gas markets and which inspire our econometric approach which is exposed in section 1.3. In section 1.4, we present the data along with some preliminary analysis of the series. The empirical findings are provided in section 1.5. Finally, section 1.6 concludes.

1.2 Distinctive features in modeling and forecasting short-term natural gas demand

At least three distinctive features are worth considering 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. In contrast, natural gas is either directly consumed by end-users (e.g., by households or industries to produce heat) or is converted into 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 for natural gas. Hence, the present analysis suggests considering two variables to predict the next day's load: the day-ahead price of natural gas along with the day-ahead spark ratio. In line with 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 in this case. Overall, that discussion suggests opting for a modeling approach that can incorporate nonlinearities 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. Moreover, 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 (Vany and Walls, 1993; Serletis and Herbert, 1999; Renou-Maissant, 2012; Thoenes, 2014) 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

[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).

spurious results. To overcome that problem, we consider the ARDL model suggested in Pesaran and Shin (1999) 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 of 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 newly formulated in this paper. We provide evidence that our data series either have a unit root or are stationary, thereby duly justifying the use of an ARDL specification.

Overall, as will be shown below, our empirical results confirm that accounting for these three features is highly consequential 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.

1.3 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 peak-load block of day t, and M_i, with $i = 1 \cdots 11$, 11 monthly dummy variables. Hence, our model has the following form: $q_t = f\left(p_t, s_t, \mathrm{M}_{i,t}\right)$.

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

[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.

1.3.1 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 those 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:

$$\begin{aligned}\Delta q_t = \alpha + \rho q_{t-1} + \theta_1 s_{t-1} + \theta_2 p_{t-1} + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} + \sum_{i=0}^{q_1-1} \gamma_i \Delta s_{t-i} \\ + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i} + \sum_{i=1}^{11} \kappa_i \mathrm{M}_{i,t} + \sum_{i=1}^{11} \zeta_i \Delta \mathrm{M}_{i,t} + \eta_t\end{aligned} \tag{1.1}$$

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_i and ζ_i represent the long- and short-run effect of the monthly dummy variables; and η_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.

As we suspect the possible presence between the price and consumption series, the results obtained from an OLS estimation of eq. (1.1) can yield inaccurate standard deviations of the

[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.

[10] By gathering the terms in levels, one can use that definition of the multipliers to express the specification in the usual ECM form: $\Delta q_t = \alpha + \rho \left(q_{t-1} - \beta_1 s_{t-1} - \beta_2 p_{t-1} - \sum_{i=1}^{11} \beta_{3,i} \mathrm{M}_{i,t} \right) + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} + \sum_{i=0}^{q_1-1} \gamma_i \Delta s_{t-i} + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i} + \sum_{i=1}^{11} \zeta_i \Delta \mathrm{M}_{i,t} + \eta_t$, where $\beta_{3,i} := -\kappa_i/\rho$ is the long-run multiplier associated with the ith monthly dummies.

estimated parameters (Pesaran and Shin, 1999). To correct for that, we follow the discussion in Section 3 of Pesaran and Shin (1999) and estimate this ARDL model using the delta-method (Δ-method hereafter). 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.

By construction, the long-run marginal impact of the electricity price to natural gas demand is β_1, which is expected to be positive, as one can presume both energies will be substitutable. Also, the long-run marginal impact of natural gas price on natural gas demand is $\beta_2 - \beta_1$, which is expected to be negative.[11]

1.3.2 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

[11] Here, the long-run multipliers represent the percentage change in the quantity demanded in the long-run resulting from a *ceteris paribus* 1 percent change in the change of the other explanatory variable (e.g., the spark ratio). Following a remark raised by two of the referees, we agree that to derive elasticity estimates, it is necessary to rely on a full-fledged structural econometric model. Though the interpretation of ARDL estimates as elasticities is quite common in the applied econometric literature in the context of modeling energy demand (Adeyemi and Hunt, 2014; Dergiades and Tsoulfidis, 2008; Hunt et al., 2003) here we abstract from such an interpretation.

asymmetric. To overcome that limitation, 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.[12] We thus decompose the spark ratio s_t as follows:

$$s_t = s_0 + s_t^+ + s_t^- \tag{1.2}$$

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:

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

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

This decomposition is then introduced in (1.1) 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:

$$\begin{aligned}\Delta q_t = \alpha + \rho q_{t-1} + \left(\theta_1^+ s_{t-1}^+ + \theta_1^- s_{t-1}^-\right) + \theta_2 p_{t-1}^G + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} \\ + \sum_{i=0}^{q_1-1} \left(\gamma_i^+ \Delta s_{t-i}^+ + \gamma_i^- \Delta s_{t-i}^-\right) + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i}^G + \sum_{i=1}^{11} \kappa_i \mathrm{M}_{i,t} + \sum_{i=1}^{11} \zeta_i \Delta \mathrm{M}_{i,t} + \eta_t\end{aligned} \tag{1.3}$$

[12] 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.

where θ_1^+ and θ_1^- are the long-run asymmetric coefficients and γ_i^+ and γ_i^- are the short-run asymmetric coefficients 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 $H_0 : \left(\rho = \theta_1^+ = \theta_1^- = \theta_2 = 0\right)$.

By construction, the NARDL specification nests three special cases: (i) a symmetric long-run relationship, that is, the null hypothesis $H_0 : \left(\theta_1^+ = \theta_1^-\right)$; (ii) a symmetric short-run relationship, that is: $H_0 : \left(\gamma_i^+ = \gamma_i^-, \forall i \in \{1, ..., (q_1 - 1)\}\right)$; 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.

1.3.3 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,[13] one can question its relevance for the present application. Arguably, the technological considerations in section 1.2 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

[13]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

regimes.[14] 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 present case, it would be preferable to examine whether an observed variation in the spark ratio does or does not have 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:

$$s_t = s'_0 + {s_t}^{>Th} + {s_t}^{\leq Th} \tag{1.4}$$

$${s_t}^{>Th} = \sum_{j=1}^{t} \Delta {s_j}^{>Th} = \sum_{j=1}^{t} \Delta s_j \mathbb{I}_{s_j > Th} \tag{1.5}$$

$${s_t}^{\leq Th} = \sum_{j=1}^{t} \Delta {s_j}^{\leq Th} = \sum_{j=1}^{t} \Delta s_j \left(1 - \mathbb{I}_{s_j > Th}\right) \tag{1.6}$$

there 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:

$$\begin{aligned} \Delta q_t = \alpha + \rho q_{t-1} + \left({\theta_1}^{>Th} s_{t-1}^{>Th} + {\theta_1}^{\leq Th} s_{t-1}^{\leq Th}\right) + \theta_2 p_{t-1}^G + \sum_{i=1}^{p} \phi_i \Delta q_{t-i} \\ + \sum_{i=0}^{q_1-1} \left({\gamma_i}^{>Th} \Delta s_{t-i}^{>Th} + {\gamma_i}^{\leq Th} \Delta s_{t-i}^{\leq Th}\right) + \sum_{i=0}^{q_2-1} \delta_i \Delta p_{t-i}^G + \sum_{i=1}^{11} \kappa_i \mathrm{M}_{i,t} + \sum_{i=1}^{11} \zeta_i \Delta \mathrm{M}_{i,t} + \eta_t \end{aligned} \tag{1.7}$$

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 ${\beta_1}^{\leq Th} := -{\theta_1}^{\leq Th}/\rho$.

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

For a given value of the threshold parameter Th, that specification can be estimated using the Δ-method. 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. That said, this procedure provides only one value of Th which is an uncertain parameter. So, we use a block bootstrapping approach to obtain both the estimator of Th and the associated standard deviation. Because of the nuisance parameter associated with the threshold value, it is important to keep in mind that the validity of the Wald-test procedure can be jeopardized (Hansen, 1996). In the present paper, the long-run and short-run symmetry restrictions are thus tested using the expF statistic described in (Andrews and Ploberger, 1994).

1.4 Preliminary Analysis of the Data

1.4.1 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 February 2, 2018. The starting date is the day on 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 the 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 1.1. 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 peak-load block that covers the hours from 9am to 8pm on the next working day.[15] 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 (Ramanathan et al., 1997; Karakatsani and Bunn, 2008; Bordignon et al., 2013), we concentrate our attention on the working days and remove the weekends and the bank holidays from the data. Hence, the day-ahead gas prices formed on a Friday value the gas to be delivered on the following Monday.

The whole dataset has a total of 625 observations. In the sequel, that data set is further divided into two parts. The first part, covering the period April 1, 2015, to December 31, 2016 (i.e., 396 observations), is only used for model estimation. The remaining part comprises 229 observations and is used for evaluating out-of-sample forecasts.

[15] Hence, the prices are formed during the day preceding the delivery (and thus the consumption) of natural gas.

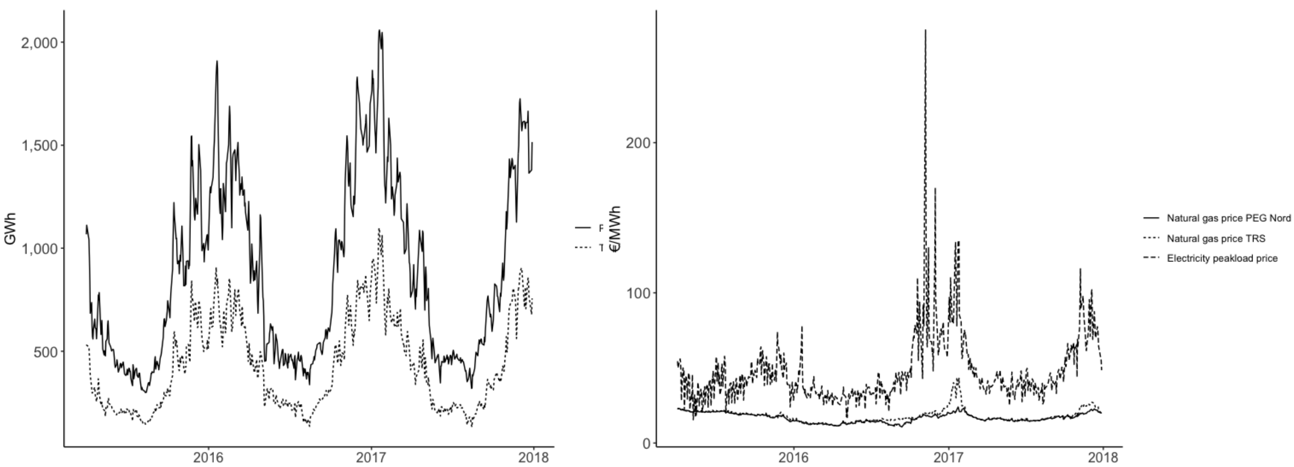
Table 1.1: Data sources.

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 (Teréga)	Consumption South	GWh
	Day-ahead price	Bloomberg	TRS, End of Day	€/MWh
Electricity	Day-ahead price	Bloomberg	Powernext Peak-load, End of Day	€/MWh

1.4.2 Descriptive statistics

Figure 1.1 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 PEG Nord and TRS demand series present a smooth yearly seasonality linked to the gradual variation in weather conditions and the use of natural gas for space heating.

Figure 1.1: Daily consumption of natural gas (left) and day-ahead energy prices in France (right).



The price plots show that the day-ahead price of electricity measured during the peak-load 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 winter 2016–2017).

Table 1.9 (in Appendix A) summarizes the descriptive statistics for the series in levels during the estimation period. The average consumption figures indicate that PEG Nord is roughly twice larger than TRS (i.e., 814.8 GWh compared with 401.8 GWh) 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. 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 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 (see the highly significant Jarque-Bera statistics in Table 1.9 (Appendix A). 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.

1.4.3 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.[16] Here, the integration properties of the data are investigated using two standard unit root tests – namely, the Augmented Dickey-Fuller (ADF) and the Phillips and Perron (PP) tests – and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for the presence of stationarity.

From the results presented in Table 1.2, it is found that none of the variables are integrated in order two which allows us to use an ARDL approach to investigate the presence of

[16] Recall that the natural gas consumption series exhibit a smooth yearly seasonality and that the presence of that seasonal pattern may bias the testing procedure used to detect the presence of a unit root. To correct for the effects of that seasonal pattern, the test results reported in Table 1.2 have been obtained with the deseasonalized consumption series resulting from a regression with monthly dummy variables.

cointegration. For the electricity and gas price series, the findings are consistent with the characteristics highlighted in earlier studies evoked in the introduction: the spot price of electricity is found to be mean reverting and the natural gas prices are non-stationary. Unsurprisingly, the properties of the spark ratio series, which is defined as the quotient of these two price series, is not straightforward to interpret. For PEG Nord, the KPSS test suggests that we are dealing with an $I(1)$ variable whereas the ADF and PP unit roots tests firmly reject the presence of a unit root. A similar issue occurs when examining the results concerning the daily consumption of natural gas. The KPSS tests strongly reject the stationarity hypotheses, whereas the ADF and PP test results suggest that the series are $I(0)$. These results may leave us wondering whether these series are stationary or not. Facing such a dilemma, we follow Cochrane (1991) who argues that the important question is not whether the time series should be unequivocally classified in the unit root or the stationary categories but, rather, how to develop the appropriate inferential procedure for the data being analyzed. In the present case, there is strong evidence that the first-differenced consumption series are stationary. Therefore, even if the support in favor of a unit root is not strong, modeling the consumption data using first-differenced series reduces the risk of spurious regression. As such, we follow Hendry and Juselius (2000) who note that *"even though a variable is stationary, but with root close to unity [...] it is often a good idea to act as if there are unit roots to obtain robust statistical inference."*(p. 21).

1.5 Empirical Findings

1.5.1 Estimation results

The lag structure is determined using the Bayesian Information Criterion (BIC). For each specification, we allow a maximum lag-length of 20 for each variable and then compare the BIC scores obtained for each possible combination of lagged values. For each model, the selected specification is the one that provides the lowest BIC score. With that selection

Table 1.2: Unit root tests.

	Levels			First differences		
	ADF	PP	KPSS	ADF	PP	KPSS
PEG Nord						
q_t	-8.41***	-7.77***	0.74_C***	-19.77***	-38.33***	0.06_C
s_t	-3.72_C***	$-12.20_{C\&T}$***	$0.27_{C\&T}$ ***	-13.81***	-69.38***	0.06_C
p_t^G	-2.09_C	-2.18_C	$0.45_{C\&T}$***	-29.42***	-29.56***	0.23_C
TRS						
q_t	-3.62***	-2.94***	2.38_C***	-16.27***	-33.52***	0.09_C
s_t	$-4.69_{C\&T}$***	$-12.66\ _{C\&T}$***	$0.19_{C\&T}$ *	-13.51***	-66.33***	0.07_C
p_t^G	-2.19_C	-2.32_C	$0.36_{C\&T}$***	-30.88***	-31.43***	0.19_C
Electricity						
p_t^E	-3.47_C***	$-9.34_{C\&T}$***	$0.11_{C\&T}$	-13.40***	-59.46***	0.08_C

Notes: For the ADF test, we apply the lag structure suggested by the Schwartz information criterion. For the PP test, the truncation lags are decided by Newey-West default. For the KPSS test, the bandwidth is selected by the Newey-West automatic selection procedure using the Bartlett kernel. 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 at 0.05** and 0.01*** significance levels, respectively.

procedure, the obtained optimal lag structure is identical for the ARDL, NARDL and TARDL models, that is: $p = 1$, $q_1 = 1$ and $q_2 = 1$, which indicates that the short-run dynamics solely consider the instantaneous impacts that the gas price and the spark ratio have on the contemporary demand changes.

For the TARDL model, the block bootstrap procedure finds that, on average, the sum of squared residuals is minimized for $Th = 0.60$ (s.d.= 0.08) in PEG Nord and $Th = 0.59$ (s.d.= 0.01) in TRS.[17] Of particular interest is the relationship 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 1.82 (respectively 1.80) 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 $[50\%, 62\%]$ range.

We now examine whether the presence of short- and long-run asymmetries is supported

[17] These values have been obtained using a time series block bootstrap procedure that comprises a total of 1,000 replications randomly drawn from the original data set using a 20-day time window.

by the data or not. From the Wald statistics $W_{SR,LR}$ and the expF statistics $expF_{SR,LR}$ reported in Table 1.3, 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 : $\left(\theta_1^+ = \theta_1^-\right)$ and $\left(\gamma_0^+ = \gamma_0^-\right)$ 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 (particularly for PEG Nord). Therefore, we further explore whether partial restrictions in the form of either short- or long-run symmetry are supported by the data or not.

The test results in Table 1.3 convey similar findings for the two nonlinear models and the two markets. We observe low values for both the Wald statistics W_{LR} and the expF statistics $expF_{SR}$ 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} and $expF_{SR}$ 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 time series diagnostic tests (see Appendix B). The test results indicate that the models are properly specified. In particular, we find no indication of serial correlation in the residuals. This finding is important for the validity of our estimates. In case of autocorrelated residuals, one could consider the approach in Lewbel and Ng (2005) that deals with the issue of nonstationarity (and nonlinearities).[18]

The estimation results obtained with our preferred specifications are reported in Table 1.4 for PEG Nord and in Table 1.5 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 derived. The first series of find-

[18]In addition, following a referee's comment, we also tested the usefulness of including a proxy of economic activity, as in Agnolucci et al. (2017), but the latter was unsuccessful in fitting the data better.

Table 1.3: Asymmetric tests

	PEG Nord		TRS	
	stat.	p-value	stat.	p-value
NARDL				
$W_{SR,LR}$	12.26	(0.06)*	13.90	(<0.01)***
W_{LR}	0.02	(0.88)	1.13	(0.28)
W_{SR}	14.84	(<0.01)***	22.28	(<0.01)***
TARDL				
$expF_{SR,LR}$	5.32	(0.09)*	7.39	(0.04)**
$expF_{LR}$	44.77	(1)	38.61	(0.37)
$expF_{SR}$	6.94	(<0.01)***	5.60	(0.02)**

Notes: The statistics $W_{SR,LR}$ and $expF_{SR,LR}$ test the null hypothesis of a fully symmetric ARDL model. The statistics W_{LR} and $expF_{LR}$ are 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 statistics W_{SR} and $expF_{SR}$ are 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.

ings concern 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 negative (as expected) for all models.

Our second series of findings focuses on the estimated long-run multipliers. In both markets, the multipliers of the spark ratio variable are, as expected, positive, and their estimated values are statistically significant at the 1% level. Regarding the gas price coefficient (that is, β_2), it is important to keep in mind that this multiplier does not fully capture the long-run influence of the price of natural gas on daily gas demand because that variable is also present in the spark ratio. Hence, the long-run multiplier that captures the long-run reaction of daily gas demand to the price of natural gas is given by the difference $\beta_2 - \beta_1$. We report these

Table 1.4: 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	5.42	(0.54)	10.03***	5.34	(0.52)	10.29***
$Coint_{t-1}$	-0.28	(0.03)	-10.03***	-0.28	(0.02)	-10.29***
Δq_{t-1}	0.08	(0.04)	2.01*	0.09	(0.04)	2.24*
Δs^{+}	0.16	(0.02)	9.18 ***			
Δs^{-}	0.06	(0.05)	1.23***			
$\Delta {s_t}^{>Th}$				0.17	(0.02)	8.45***
$\Delta {s_t}^{\leq Th}$				0.11	(0.03)	4.20***
Δp^{G}	0.15	(0.13)	1.14	0.16	(0.13)	1.22
ΔM$_1$:Jan	0.23	(0.07)	3.35***	0.25	(0.07)	3.62***
ΔM$_2$:Feb	0.06	(0.09)	0.67	0.08	(0.09)	0.87
ΔM$_3$:Mar	0.09	(0.10)	0.82	0.10	(0.10)	1.00
ΔM$_4$:Apr	0.07	(0.10)	0.61	0.08	(0.11)	0.79
ΔM$_5$:May	-0.18	(0.10)	-1.69*	-0.15	(0.10)	-1.47
ΔM$_6$:Jun	-0.17	(0.10)	-1.63*	-0.15	(0.10)	-1.48
ΔM$_7$:Jul	-0.17	(0.09)	-1.68*	-0.16	(0.09)	-1.65*
ΔM$_8$:Au	-0.33	(0.09)	-3.63***	-0.32	(0.09)	-3.60***
ΔM$_9$:Sep	-0.25	(0.08)	-3.00***	-0.24	(0.08)	-3.00***
ΔM$_{10}$:Oct	-0.12	(0.07)	-1.70	-0.12	(0.07)	-1.67*
ΔM$_{11}$:Nov	-0.04	(0.05)	-0.75	-0.04	(0.05)	-0.79
Long-run multipliers						
s_t	0.43	(0.02)	29.30***	0.54	(0.01)	39.9***
p_t^G	0.34	(0.07)	4.74***	0.45	(0.07)	8.02***
M$_1$:Jan	0.15	(0.07)	2.09*	0.24	(0.06)	3.78***
M$_2$:Feb	0.26	(0.07)	3.02***	0.36	(0.06)	5.66***
M$_3$:Mar	0.15	(0.06)	2.40*	0.27	(0.06)	4.62***
M$_4$:Apr	-0.18	(0.05)	-3.63***	-0.17	(0.04)	-3.76***
M$_5$:May	-0.61	(0.05)	-12.63***	-0.49	(0.04)	-11.48***
M$_6$:Jun	-0.87	(0.05)	-18.6***	-0.79	(0.04)	-18.91***
M$_7$:Jul	-0.99	(0.05)	-20.68***	-0.94	(0.04)	-21.83***
M$_8$:Aug	-1.05	(0.05)	-22.74***	-0.99	(0.04)	-23.92***
M$_9$:Sep	-0.79	(0.05)	-16.13***	-0.75	(0.04)	-17.27***
M$_{10}$:Oct	-0.35	(0.05)	-6.95 ***	-0.34	(0.06)	-7.52***
M$_{11}$:Nov	-0.17	(0.05)	-3.40***	-0.16	(0.04)	-3.58***
Bounds Test	F-statistic	$I(0)$	$I(1)$	F-statistic	$I(0)$	$I(1)$
W_{coint}	10.12***	4.29	5.61	12.13***	4.29	5.61
$Adjust\ R^2$	0.95			0.98		
Observations	395			395		

Notes: The table reports the estimation results obtained with the Δ-method for PEG Nord for the period April 1, 2015, to December 31, 2016. 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.

Table 1.5: 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						
(Constant)	6.44	(0.51)	12.53***	6.73	(0.55)	12.06***
$Coint_{t-1}$	-0.34	(0.02)	-12.53***	-0.34	(0.03)	-12.06***
Δq_{t-1}	0.09	(0.04)	2.20**	0.12	(0.04)	2.81**
Δs^{+}	0.15	(0.02)	8.42***			
Δs^{-}	0.07	(0.05)	1.54			
${\Delta s_t}^{>Th}$				0.15	(0.02)	6.96***
${\Delta s_t}^{\leq Th}$				0.11	(0.02)	4.56***
Δp_t^G	0.32	(0.13)	2.40**	0.30	(0.13)	2.18**
ΔM$_1$:Jan	0.18	(0.07)	2.57***	0.17	(0.07)	2.34**
ΔM$_2$:Feb	0.04	(0.09)	0.41	0.014	(0.09)	0.15
ΔM$_3$:Mar	0.01	(0.10)	0.04	-0.02	(0.11)	-0.22
ΔM$_4$:Apr	0.10	(0.11)	0.92	0.06	(0.11)	0.59
ΔM$_5$:May	-0.03	(0.10)	-0.30	-0.06	(0.11)	-0.59
ΔM$_6$:Jun	-0.04	(0.10)	-0.35	-0.06	(0.11)	-0.59
ΔM$_7$:Jul	-0.02	(0.10)	-0.20*	-0.05	(0.10)	-0.48
ΔM$_8$:Au	-0.23	(0.09)	-2.42*	-0.24	(0.09)	-2.56**
ΔM$_9$:Sep	-0.17	(0.08)	-1.98*	-0.18	(0.09)	-2.13**
ΔM$_{10}$:Oct	-0.08	(0.07)	-1.09	-0.09	(0.07)	-1.28
ΔM$_{11}$:Nov	-0.08	(0.05)	-1.68*	-0.10	(0.05)	-1.98*
Long-run multipliers						
s_t	0.51	(0.02)	34.24***	0.51	(0.01)	40.90***
p^G	0.33	(0.07)	4.39***	0.13	(0.07)	1.83*
M$_1$:Jan	0.09	(0.07)	1.37	0.06	(1.00)	0.31
M$_2$:Feb	0.23	(0.07)	3.50***	0.17	(0.06)	2.68***
M$_3$:Mar	0.09	(0.06)	1.46	0.04	(0.06)	0.70
M$_4$:Apr	-0.20	(0.05)	-4.31***	-0.39	(0.04)	-8.43***
M$_5$:May	-0.58	(0.05)	-12.87***	-0.65	(0.04)	-14.66***
M$_6$:Jun	-0.84	(0.04)	-19.08***	-0.87	(0.04)	-20.51***
M$_7$:Jul	-0.99	(0.05)	-21.81***	-1.01	(0.04)	-23.03***
M$_8$:Aug	-1.12	(0.04)	-25.74***	-1.14	(0.04)	-27.11***
M$_9$:Sep	-0.81	(0.05)	-17.73***	-0.83	(0.04)	-18.82***
M$_{10}$:Oct	-0.45	(0.05)	-9.43***	-0.47	(0.04)	-9.96***
M$_{11}$:Nov	-0.23	(0.05)	-4.73***	-0.23	(0.04)	-4.83***
Bounds Test	F-statistic	$I(0)$	$I(1)$	F-statistic	$I(0)$	$I(1)$
W_{coint}	16.58***	4.29	5.61	18.46***	4.29	5.61
Adjust R^2	0.97			0.97		
Observations	395			395		

Notes: The table reports the estimation results obtained with the Δ-method for TRS for the period April 1, 2015, to December 31, 2016. 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.

differences in Table 1.6. The obtained values are negative, as expected, but have very low magnitudes (particularly for PEG Nord). This finding indicates that, in the long run, the daily gas consumption is not very sensitive to the gas price. As one may wonder whether that consumption could be completely unaffected by that price in the long run, we also report the test results for the null hypotheses $\beta_2 - \beta_1 = 0$ in Table 1.6. For each market, we found that this hypothesis is firmly rejected by the data for the NARDL and the TARDL models. Thus, in the long run, the gas price has a small but significant impact on observed consumption levels.

Table 1.6: Long-run reaction of natural gas demand to the price of natural gas.

	PEG Nord			TRS		
	value	stat.	p-value	value	stat.	p-value
NARDL						
$\epsilon^{LR}_{Gas} := (\beta_2 - \beta_1)$	-0.09			-0.18		
$H_0 : (\beta_1 - \beta_2 = 0)$		15.44	<0.01***		151.70	<0.01***
TARDL						
$\epsilon^{LR}_{Gas} := (\beta_2 - \beta_1)$	-0.09			-0.38		
$H_0 : (\beta_1 - \beta_2 = 0)$		166.68	<0.01***		174.17	<0.01***

Notes: The table reports the difference $(\beta_2 - \beta_1)$ computed from the estimation results. It also reports the result of a Wald test for NARDL model and the expF test statistic of Andrews and Ploberger (1994) for the null hypothesis of zero long-run reaction of demand to gas price. Asterisks indicate rejection of the null at 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 and that the magnitude of that impact is substantially larger when the spark ratio is larger than the threshold value. These results are fully consistent with the expectations derived from the technological considerations discussed in section 1.2. For the NARDL model, the asymmetric coefficients obtained in the two markets are positive. 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 a positive (respectively negative) long-run relationship between natural gas consumption and electricity (respectively natural gas) price and of short-run dynamics that are clearly asymmetric, as was expected from the technology considerations above.

1.5.2 Out-of-sample analysis

As the NARDL and TARDL models are not nested, it is easy to decide which model should be used using standard tests. To compare the predictive performance of these two models, we now use the estimates from our preferred NARDL and TARDL specifications to generate out-of-sample forecasts for the evaluation period.[19]

To gain insight on their predictive accuracy, we use three benchmarks. The first is given by a simple AR(1) model, which is a standard reference commonly used in the forecasting literature. The second benchmark is the linear ARDL model presented above. That benchmark should provide useful insights on the performance gains resulting from supplementing an ARDL specification with nonlinear components. Lastly, our third benchmark is formed by the next day's demand forecast 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 be readily used as a benchmark. In the TRS zone, there are two pipeline operators (GRTGaz and Terēga) 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 1.7 reports a common measure of accuracy of the predicted values of the next day's consumption: the Root Mean Square Error (RMSE) which is measured in GWh over the entire evaluation period. As the consumption of natural gas is substantially larger during the winter season than during the summer, when gas-fueled heating systems are turned off, this table

[19] Recall that this period has 229 observations, which is large enough to compare the predictions obtained with different models.

also reports the prediction errors obtained for these two specific seasons in our out-of-sample validation period.

Table 1.7: Prediction error statistics values

		NARDL	TARDL	AR(1)	ARDL	TSO
Validation period	PEG Nord	26.18	**11.77**	55.06	39.15	47.14
	TRS	23.59	**3.77**	59.70	17.01	23.76
Summer subperiod	PEG Nord	31.04	**14.75**	69.43	37.76	31.43
	TRS	24.77	**2.78**	75.14	17.08	19.94
Winter subperiod	PEG Nord	20.35	**8.20**	70.12	40.70	65.57
	TRS	10.51	**5.32**	33.07	16.93	27.59

Notes: The RMSE, measured in GWh, is successively evaluated: for the entire evaluation period (January 1, 2017, to February 2, 2018), for the "gas winter" subperiod gathering all the observations between Nov. 1 and March 31 and for the "gas summer" subperiod (i.e., all the observations between April 1 and Oct. 31). The figures in bold indicate that this model has the lowest error value among the three models.

Overall, we note that the results obtained for PEG Nord and TRS during the out-of-sample forecast validation period and during the two subperiods are qualitatively similar. In most cases, the RMSE statistics obtained for the PEG Nord market are larger than the ones obtained for the TRS market, which is consistent with the relative sizes of these two markets.

An examination of the performances of our three benchmarks conveys the following findings. Observe that, as can be expected, whatever the market and the period under scrutiny, the forecasts issued by the TSOs and the ones obtained with a linear ARDL model are substantially more accurate than the ones emanating from a simple AR(1) specification. This finding confirms that a simple autoregressive structure is not adapted to predict the next day's load. Interestingly, one can also note that, in all cases, the RMSE statistics obtained with a linear ARDL model are markedly lower than the ones computed from the forecasts published by the TSOs. Given the excellent performance of the linear ARDL model, it is not surprising to observe that the two nonlinear extensions of that model also compete very favorably with the TSO benchmark. Overall, these findings suggest that the forecasting performance of the tools currently used by infrastructure operators remains poor, which justifies the current heightened regulatory attention paid to that issue.

We now examine the performance of our preferred nonlinear models. For both markets, we observe that our TARDL model systematically provides the lowest RMSE figures and outperforms both the three benchmarks and the NARDL model. In contrast, one can note that a nonlinear specification based on a positive and negative partial sum decomposition of the spark ratio hardly surpasses the ARDL and the TSO models. Indeed, the prediction errors obtained with the NARDL model are either commensurate or larger than these two benchmarks (cf. Table 1.7: the RMSE figures obtained for PEG Nord during the summer and the ones for TRS during the entire validation period). These findings suggest that one should prefer the TARDL model to the NARDL model and thus use that model to assess the impacts that gas and electricity prices have on daily gas demand.

In order to statistically compare the predictive accuracy of our models, we also perform the test procedure proposed by Diebold and Mariano (1995) and extended in Harvey et al. (1997). The null hypothesis is that the two forecasting models have the same forecast accuracy and the alternative hypothesis is that a baseline model (labeled B) is more accurate than the reference model (labeled A).[20] A negative value for the Diebold-Mariano statistic indicates that the baseline model (model B) delivers significantly better forecasts. The Diebold-Mariano statistics reported in Table 1.8 make use of the TARDL specification as a reference model. These test results confirm the preceding findings as the TARDL model: (i) performs significantly better than the AR(1), the ARDL and the NARDL model in all cases, and (ii) is also significantly more accurate than the forecasts issued by infrastructure operators during the winter subperiod.

This excellent performance of the TARDL model emphasizes the pivotal role of the spark ratio for the purpose of forecasting natural gas demand. Our findings confirm the intuition that allowing for a possibly non-zero threshold does help in forecasting the natural gas demand as natural gas consumption unequivocally 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

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

Table 1.8: Diebold-Mariano test statistics

		NARDL	TARDL	AR(1)	ARDL	TSO
Validation period	PEG NORD	38.91***	-	12.95***	15.79***	1.31*
	TRS	28.44***	-	13.6***	21.20***	1.43*
Summer subperiod	PEG NORD	31.47***	-	9.81***	15.19***	0.73
	TRS	22.70***	-	10.37***	18.54***	1.15*
Winter subperiod	PEG NORD	27.72***	-	8.99***	7.84***	6.49***
	TRS	5.58***	-	3.47***	5.20***	7.33***

Notes: This table reports the Diebold and Mariano (1995) (hereafter DM) test of the null hypothesis of no difference in the accuracy of the compared forecasts. The TARDL forecasts are used as a baseline model. The test statistics are evaluated: for the entire evaluation period (January 1, 2017, to February 2, 2018), for the "gas winter" subperiod gathering all the observations between Nov. 1 and March 31 and for the "gas summer" subperiod (i.e., all the observations between April 1 and Oct. 31). Our implementation of that test uses a loss differential defined as the difference between squared forecast errors. Numbers are DM tests statistics. Asterisks indicate rejection of the null hypothesis at 0.10*, 0.05** and 0.01*** significance levels, respectively.

spark ratio contains much information pertaining to natural gas demand.

1.6 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 substantial enough to provide reasonably accurate predictions of 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 forecasts. That said, the fact that a relatively simple econometric model, based solely 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,in the long run, the marginal impact of natural gas price on daily gas consumption is negative and that we are dealing with a very small price-responsive demand. In the long run, the daily gas demand is also positively affected by the price of 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: Descriptive statistics

Table 1.9: 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.

Appendix B: Diagnostics tests

The models were subjected to several diagnostic tests to detect: the presence serial correlation, the presence of heteroscedasticity, and a possible functional misspecification. The test results are detailed in Table 1.10.

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 misspecification based on the powers of the fitted value of consumption shows no sign of functional misspecification for both models in each market. 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. 1.2 and Fig. 1.3). 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

Table 1.10: Diagnostics test.

	NARDL model		**TARDL model**	
	Statistic	p-value	Statistic	p-value
PEG Nord				
A: Serial correlation L.-B.(5)	1.70	0.19	1.47	0.22
B: Heteroscedasticity ARCH(5)	0.46	0.99	0.59	0.98
C: Functional form RESET(2,3)	1.26	0.26	0.87	0.35
D: KS-GMM.	-	0.22	-	0.57
TRS				
A: Serial correlation L.-B.(5)	0.53	0.46	0.24	0.62
B: Heteroscedasticity ARCH(5)	0.05	1.00	0.04	1.00
C: Functional form RESET(2,3)	1.85	0.15	1.00	0.36
D: KS-GMM	-	0.67	-	0.78

Notes: L.-B.(5) is the Ljung-Box Q-statistic for the null hypothesis of no serial correlation up to the 5th 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. KS-GMM is the Kolmogorov-Smirnov test for the null hypothesis of a Gaussian Mixture Model based on Smirnov (1948). As the KS-GMM estimation is based on a bootstrap procedure, we solely report the p-value. Asterisks indicate rejection of the null hypothesis at 0.10*, 0.05** and 0.01*** significance levels, respectively.

of the models over the estimation period. The results gained from Kolmogorov-Smirnov tests indicate that, in all cases, a Gaussian Mixture Model adequately describes the distributional properties of the residual series.[21]

[21] From a technical perspective, these distributional properties are compatible with the theroretical conditions in Pesaran and Shin (1999) for the validity of the model estimates as Pesaran and Shin (1999) simply stipulate that the error terms should be *"serially uncorrelated disturbances with zero means and constant variance-covariances."*

Figure 1.2: CUSUM & CUSUMQ test with the NARDL models for PEG Nord (on the left) and TRS (on the right)

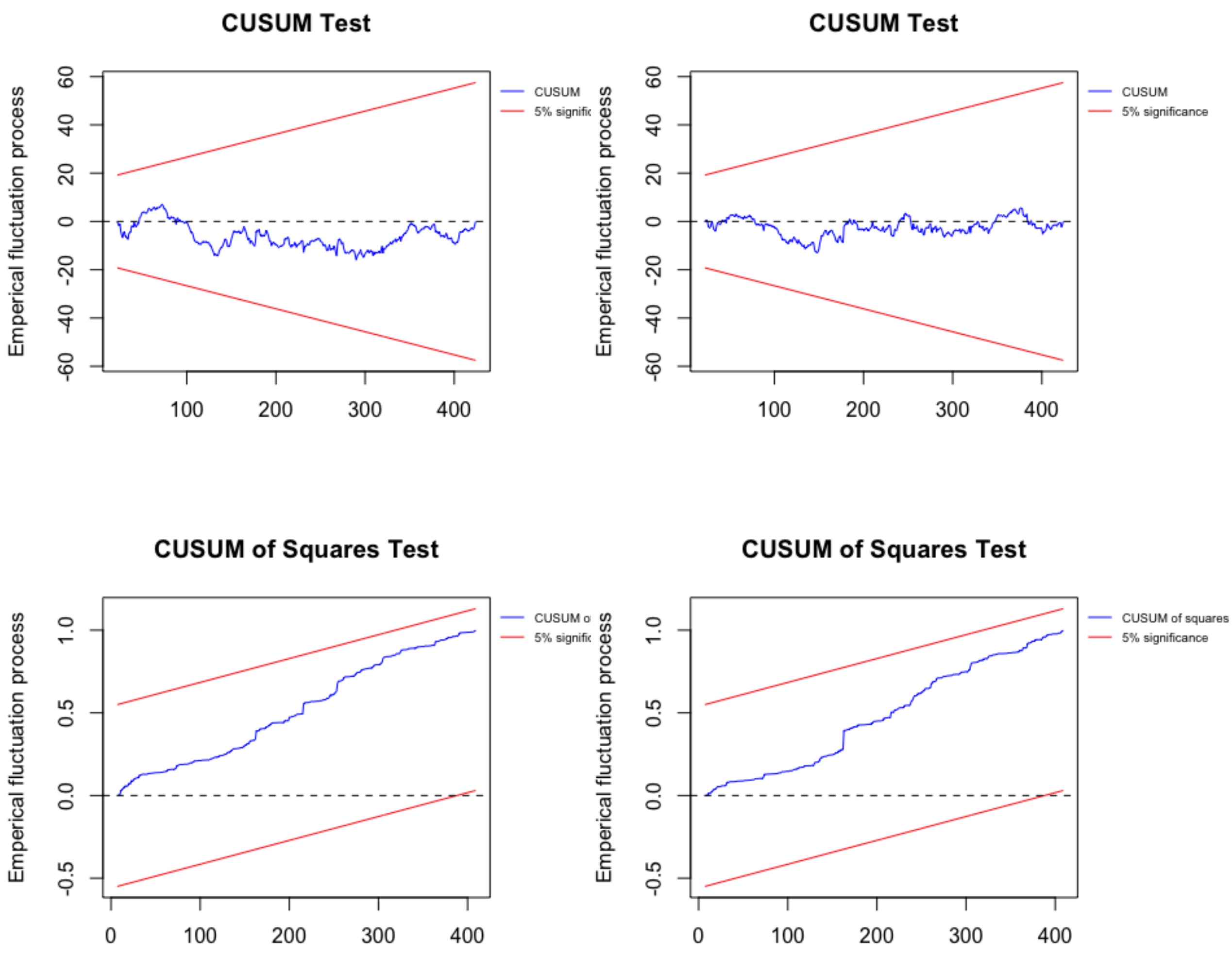
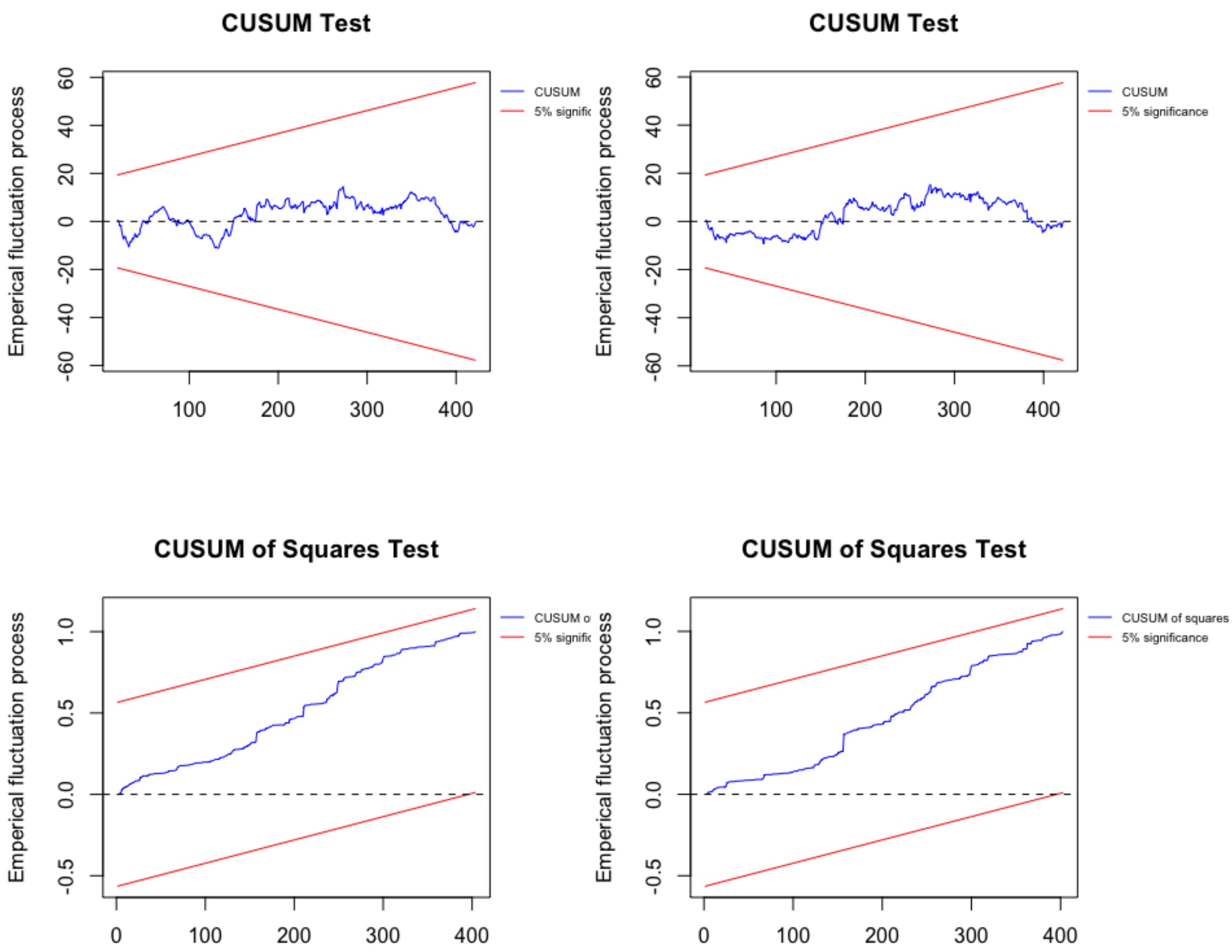


Figure 1.3: CUSUM & CUSUMQ tests with the TARDL models for PEG Nord (on the left) and TRS (on the right)



Chapter 2

Considering real-time demand to forecast the U.S. natural gas price in real-time: The role of temperature data

[This chapter is based on Moussa et al. (2020)]

2.1 Introduction

This paper develops a forecasting analysis of real natural gas prices at the Henry Hub in real-time. We make use of *state-of-the-art* Bayesian VAR econometrics allowing for stochastic volatility and/or fat tails in an environment where variable selection can take place. We show that considering a model *à la* Baumeister and Kilian (2012) for natural gas along with the price of crude oil as an additional variable significantly improves the forecast accuracy at horizons going from 3 to 12 months. The originality of our approach lies in considering directly – through the consumption of gas – the demand side as in Baumeister et al. (forthcoming) in addition of the supply side already present in Baumeister and Kilian (2012).[1] We further show

[1] Baumeister and Kilian (2012) also consider the demand side through a quite general activity index which is likely to proxy for the state of the economy. Our approach is more direct in the consumption of natural gas itself is considered to proxy the demand for gas.

that considering an additional variable such as temperature, which is available in real-time, increases forecast accuracy at all horizons. Using temperature data allows to include variations in demand in real-time in our models and is shown to be critical to the performance of our forecasts at all horizons. To our best knowledge, this is the very first time such an approach is retained to forecast commodity prices.

Natural gas is a key industry in the U.S. with a total revenue for distributors amounting to $101 bn in 2018. There are more than 75 million residential, commercial and industrial natural gas customers in the U.S. and, today, natural gas meets more than one-fourth of the United States' energy needs (American Gas Association, 2019). The U.S. natural gas price is closely watched by consumers as it directly affects their budget as well as their home heating choices. It is also a key commodity price for the U.S. economy and has an impact on inflation expectations, consumption level and consumer confidence in the economy. Therefore, it is also of vital importance for central bankers. Moreover, natural gas has a more and more prominent role in generating power.[2]

The shale gas revolution also plays a central role in shaping energy economic equilibrium in the U.S.[3] The production of natural gas in the U.S. reached a new high in 2018 and the trend is still positive (IEA, 2017). In addition, the development of shale gas extraction has an impact on external balance as the U.S. export more than they import for the first time

[2]Steam turbines (which can also be powered by oil or coal) combust fuel to generate steam, which is used in a steam turbine to generate electricity. Combined-cycle units heat up fuel and use the fuel-air mixture to spin gas turbines and generate electricity. The waste heat from the gas turbine is used to generate steam for a steam turbine that generates additional electricity. Based on the U.S. Energy Information Administration's (EIA hereafter) December 2018 monthly electric generator inventory of utility-scale generation, 31.3 gigawatts (GW) of generating capacity were added in the United States in 2018 and 18.7 GW of capacity were retired. The 2018 annual capacity additions were the largest since 48.8 GW that were added in 2003. Most of the additions happened in the second half of the year, while the retirements occurred mostly in the first half. More than 60% of electric generating capacity installed in 2018 was fueled by natural gas (EIA, 2019b) and since April 2018, the U.S. produce more power using natural gas than using coal (EIA, 2019a). The amount of generating capacity from natural gas-fired combined-cycle (NGCC) plants has grown steadily over time, and in 2018, surpassed coal-fired plants as the technology with the most electricity generating capacity in the United States. As of January 2019, U.S. generating capacity at NGCC power plants totaled 264 gigawatts (GW), compared with 243 GW at coal-fired power plants. Total capacity for generating power in the United States across all types of natural gas-fired generating technologies surpassed coal as the primary capacity resource more than 15 years ago.

[3]See Hausman and Kellogg (2015) about the welfare and distributional implications of the U.S. shale gas revolution.

in 2017 (EIA, 2018). Some authors, however, explain in which respect the shale oil and gas revolutions have contrasted impacts because of transport constraints inherent to gas.[4]

Investments in the gas industry are significant and often non-reversible. Having better price visibility is important for all players in the gas industry. Natural gas infrastructure operators (TSO) can establish the transportation cost which is necessary to operate, balance and maintain the pipeline networks in order to ensure the security of supply. Power generators rely on forecasts to compete more efficiently into the natural gas market, especially in a world where the role of natural gas is changing. Indeed, because of the rise of intermittent renewable energy sources of electricity, natural gas-based thermal generation is increasingly used as a back-up technology. For reasons above, providing better-quality long-term forecasts allows market participants to improve visibility into the evolution of their investments and are, more generally, critical for all participants to the natural gas market.

There are quite a large number of papers that have tried to forecast commodity prices. Some papers only rely on the price time series itself (Wang et al. (2020), Gao et al. (2020)) while others make use of the information in a few fundamental variables such as financial variables as equity and bond yields (Bessembinder and Chan (1992), Baumeister et al. (2015)), exchange rates (Chen et al. (2010)), futures prices (Alquist and Kilian (2010b) and Baumeister et al. (2018)) or open interest in futures markets (Hong and Yogo (2012)). A quite different approach was developed for oil prices in Alquist et al. (2013) and Baumeister and Kilian (2012) and aims at modeling the oil market in a vector autoregression (VAR) framework where the variables at stake are commonly the oil price, the oil production, oil inventories and an index of

[4] Kilian (2016) writes: "This conclusion stands in marked contrast to the natural gas sector. Although at first glance the shale oil revolution appears to be closely related to the shale gas revolution discussed in Hausman and Kellogg (2015), there is actually an important difference. Unlike in the gasoline and diesel market, there has been a noticeable decline in the U.S. wellhead price of natural gas since 2008 relative to natural gas prices abroad (notwithstanding a partial reversal in recent years). Access to inexpensive natural gas benefits the petrochemical industry, for example. U.S. natural gas prices have been low because the natural gas market has never been a global market, but a regional market. Natural gas is transported by pipeline. Although natural gas may be cooled down and liquefied, allowing it to be shipped as liquefied natural gas (LNG) to any port in the world, the cost of LNG shipping is high and the infrastructure required to load and unload LNG is expensive. This fact has prevented the integration of regional natural gas markets and the emergence of a global price thus far. This means that, for the time being, the price of U.S. natural gas has been determined by domestic demand rather that global demand, allowing for a greater price response to increased domestic supply."

economic activity. This method delivered very promising forecasts and was extended in various directions both considering methodological and data improvements. In particular, the recent paper by Baumeister et al. (forthcoming) focuses on enhancing the economic activity index used in Alquist et al. (2013) and Baumeister and Kilian (2012) – and which is a proxy for oil demand – as well as developing a consumption-based VAR model which replaces the oil production with the petroleum consumption on the ground that "fluctuations in the demand for refined products will translate into changes in the demand for crude oil and thus have predictive power for the future path of the real price of crude oil." (Baumeister et al. (forthcoming), p. 11). This consumption-based model leads to better forecasts in most cases thereby emphasizing the central role of demand in the oil price dynamics.

Our approach is in the vein of the original contribution of Alquist et al. (2013) and Baumeister and Kilian (2012) but with some essential specificities related to natural gas. We estimate a four-variable VAR model including the U.S. real price of natural gas (Henry Hub), the U.S. natural gas production, the U.S industrial production and the real price of crude oil.[5] In so doing, our model follows Nguyen and Okimoto (2019). While it is common to include energy-commodity production in the VAR to consider the supply side, the use of the U.S. industrial production is less common and is motivated, in our framework, by the regional feature of the gas market.[6]

Our motivation for including the price of oil comes from a long tradition in energy economics which considers the strong link between these two energy sources (see Bachmeier and Griffin (2006), Brown and Yücel (2008) and more recently, Bello et al. (2017), Bunn et al. (2017) and Jadidzadeh and Serletis (2017)) while such a relationship has been called into question in recent years following the shale gas revolution (Caporin and Fontini (2017)). The intuition behind this choice is that oil and gas are partly substitutable and are extracted in the same wells. Our results show that, in contrast to Hou and Nguyen (2018), where oil price shocks

[5] Our sample periods covers the last two decades (1997-2018) and the major reforms introduced in 1989 (Natural Gas Wellhead Decontrol Act) are not in our sample period. The major reforms motivate the Markov switching approach in Hou and Nguyen (2018), who covers the 1980-2016 period.

[6] Any attempt to include a world economic activity index in our empirical framework were, in all cases, not fruitful.

are shown to be of minor importance to modeling U.S. gas prices, oil prices are very helpful in our VAR framework. Indeed, models which does not include crude oil prices performs poorly with respect both to the no-change and have more generally low predictive power.

Beyond this standard VAR model whose aim is primarily to consider supply and demand in an endogenous way, we further take into account the role of temperature in light of its critical role in shaping natural gas consumption. This fact is well established since at least Berrisford (1965) and has been studied in numerous papers since then (Timmer and Lamb (2007) is an emblematic analysis of such a strong relationship for the U.S.). As temperature is, by nature, an exogenous variable over reasonable period of time and is available in real-time, it is a perfect natural candidate to forecast gas prices.[7] Including temperature in our VAR permits to reinforce the consumption-based feature of our model in line with Baumeister et al. (forthcoming) but keeping in mind that, in contrast to variables aiming to proxy demand in the VAR price forecasting literature, temperature has the exceptional characteristic to be a truly available in real-time which is an essential feature in our setting.

Our contribution to the literature is threefold. First and foremost, this paper is the first to provide monthly real-time forecasts of the real price of natural gas for the U.S. following the tradition initiated in Baumeister and Kilian (2012) who, as we reiterate, take advantage of the data that are freely available in the *Energy Monthly Review* on the EIA website to construct a database in real-time.[8] As such, our paper is likely to serve as a benchmark for future research on gas markets and, as pointed out above, will be useful for market participants, policy makers and regulators as well as for central bankers. Second, we provide evidence that our Bayesian

[7]In contrast with Huurman et al. (2012), who focus on power prices, we do not forecast the weather to forecast gas prices as our predictions are over longer term (1 to 12 months). Moreover, note that only a few papers in the field of energy economics have considered temperature so far (see Panagiotidis and Rutledge (2007) among others).

[8]We are only aware of the paper by Ferrari et al. (2019) who provide energy prices forecasts, including those for natural gas, at the quarterly frequency. Their empirical approach is very different from ours and relies on a global dataset including a large number of macroeconomic and financial variables. Other important contributions which dedicated to forecasting non-oil energy prices include Baumeister et al. (2017a) who forecast the gasoline price in the U.S. To the best of our knowledge, there is no empirical study aiming at providing natural gas price forecasts in an environment considering fundamental variables of the natural gas market. The recent contribution by Thomas (2020) is noteworthy as it provides evidence that a very parsimonious model is able to deliver highly competitive forecasts using the convenience yield to consider expectations.

VAR models produce promising forecasts of the real natural gas price at horizons from 1 to 12 months with, for instance, some improvements relative to the no-change forecasts up to 47% at the horizon of 12 months.[9] These massive improvements highlight the crucial importance of fundamental in modeling the natural gas market dynamics. Third, we provide evidence of the central role of considering demand in real-time through temperature data thereby avoiding the issue of using revised data for estimation. While such an approach is not feasible in the case of the world oil market, it is particularly relevant in the case of the U.S. natural gas which is a very large but regional market. Models including temperature as an exogenous variable deliver superior forecasts in particular at longer horizons.

The rest of the paper is as follows. The next section makes an exposition of the data and how the data is treated before estimation. Models are presented in Section 2.3 and estimation algorithms and forecasting procedure are discussed in Section 2.4. Section 2.5 presents the empirical results and evaluates the accuracy of the various forecasts. Finally, Section 2.6 provides concluding remarks and possible avenues for future research.

2.2 Data

The VAR models employed in our study contain four endogenous variables. As main drivers of the real gas price, we considered both supply and demand sides, namely, the log-difference in US gas production and the log in US gas consumption. In addition, we include the real crude oil price for two reasons. First, oil price can be considered as a proxy for the future state of the US economy and, as such, it further includes a demand component in our econometric specification. Second, as gas prices are quite often oil price indexed, oil price can be regarded as a good predictor for the future gas price. The choice of these variables is relatively standard and follows the existing literature.[10]. Our choice to proxy for the crude oil price goes to the spot West Texas Intermediate oil price (WTI) which is available in real-time and is very similar

[9] As will be clear below, such an improvement is expressed in terms of the ratio of the MSPE for the model at stake relative to the MSPE of the no-change forecast.

[10] Nguyen and Okimoto (2019), Hou and Nguyen (2018), Pindyck (2004), Brown and Yücel (2008), and Jadidzadeh and Serletis (2017) among many others.

to other existing benchmarks such as the RAC. The WTI price is know as a world benchmark and is used often used in academic studies making use of oil price. In addition, we include the temperature as exogenous variable. This is motivated by the fact that gas demand, from both industrial and residential heating sectors, is highly sensitive to temperature (Müller et al. (2015)). We return to the detail of the construction of this variable below.

2.2.1 Real-time data sources

As it is well documented in Baumeister and Kilian (2012), using real-time data markedly improves forecast accuracy. Some of data used in our work are available in real-time and are not subject to revision over time, like the WTI and the Henry Hub natural gas price (HH). For the remainder, however, we need to construct the real-time data set extracted from various sources. Therefore, our first contribution in this study is to construct a real-time data set for the U.S. gas market as done originally in Baumeister and Kilian (2012). Gas production and consumption were hand collected from the *Monthly Energy Review* published by the U.S. Energy Information Administration (EIA).[11] Monthly averages of the daily WTI spot price and Henry Hub gas prices were obtained from the FRED database.[12] The dataset is monthly and runs from January 1997 to August 2018, as Henry Hub data is available from January 1997 onward. As our focus in this paper is to forecast real gas price, nominal oil and gas prices were therefore deflated by the real-time U.S. consumer price index for all urban consumers.[13]

Much of the existing studies on forecasting of oil prices using VAR models uses inventories as a proxy for precautionary demand (Baumeister and Kilian, 2012; Kilian and Murphy, 2014b; Baumeister and Hamilton, 2019b). However, storage operators exploit the highly predictable

[11] For checking purpose, digitization of the *Monthly Energy Review* is realized and the collected data are compared with the hand collected series.

[12] `https://fred.stlouisfed.org/`

[13] Real-time data for the monthly seasonally adjusted U.S. consumer price index for all urban consumers are obtained from the Economic Indicators published by the Council of Economic Advisers. These data are made available by the FRASER database of the Federal Reserve Bank of St. Louis (`https://fraser.stlouisfed.org/theme/economic-data`). Additional real-time consumer price index (CPI) data were obtained from the macroeconomic real-time database of the Federal Reserve Bank of Philadelphia (`https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-for-macroeconomists`).

seasonality in the gas market. Therefore, only unexpected shifts in gas demand, which are caused by extraordinary weather conditions, are expected to be relevant for explaining gas market price.[14]

Accordingly, we drwa our attention to deviations from the normal seasonal meteorological pattern as a determinant of gas prices. Following Müller et al. (2015), we use the normalized cumulated heating degree days (CHDD). More explicitly, we cumulate heating degree days over a winter, from which we subtract the average of cumulated heating degree days over previous winters.[15] This allow us to cover effects of unexpected temperature conditions on gas prices. Figure 2.2 in Appendix A show the original data of heating degree days and our measure of temperature.

2.2.2 Nowcasting

When using real-time data, one complication arises because data are often available with a delay, from one to three month depending on the variables. The real-time gas production becomes available at a point in time, with a lag of 3 months for all the sample between January 1997 and August 2018, while consumer price index has a delay of only one month. The gas consumption and the gas stock have the same publication lags, namely, with one month for the period between January 1997 and June 2002 and with three months from July 2002 to the end of our sample.

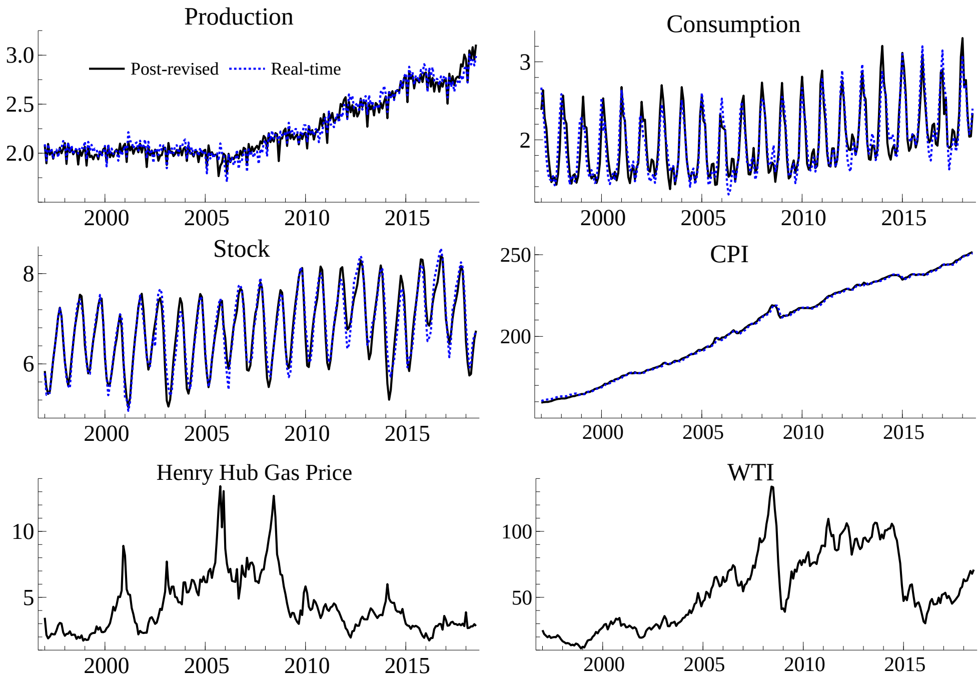
In order to obtain nowcasts for vintages of the gas production and the consumer price index, we follow Baumeister and Kilian (2012) and extrapolate data based on the average rate of change in variables up to the required date. The situation is very different for the gas consumption and inventory which require specific processing in order to take into account the marked seasonality characterizing these variables. Hence, we rely on a SARIMA specification.

[14] In preliminary experiments, we estimated our VAR models using the variation of the real-time U.S. gas stock and, in all cases, the forecasting performance was significantly lower. Results are available upon request.

[15] Our measure of the normalized cumulated heating degree days is :
$\Lambda_{d,w} = CHDD_{d,w} - \frac{1}{w-1}\sum_{l=1}^{w-1} CHDD_{d,l} \quad for \quad 1 \leq d \leq 182$, where 182 is the number of days in a winter (from the 1st of October to the 31st of March). Λ returns to zero during summer. For more details about this measure see Müller et al. (2015).

The real-time series are plotted in Figure 2.1 along with the post-revised series (sub-figures from 1 to 4). The gap between each pair of series is small on average, but is more pronounced for gas production because of the three-month delay between the first release and the last data revision.

Figure 2.1: Real-time and Post-revised data variables



Gas production, consumption and stock are expressed in Billion Cubic Feet. Henry Hub price is in Dollar per Million Btu and WTI is in Dollars per Barrel.

2.3 Models

This section introduces the specification of Bayesian VAR (BVAR) models and an overview of the estimation methods. Linear frameworks include several BVARs possibly with stochastic volatility (SV) and/or Student-t distributed shocks. We show how to extend the VAR models allowing for SV and fat tails to include some exogenous regressors. We also show how to implement variable selection for these models.

2.3.1 Bayesian VAR with stochastic volatility and fat tails

We here show how the VAR model with stochastic volatility and Student's t-distribution (tVARVOL) of Chiu et al. (2017b) can allow for including some exogenous variables (tVARVOLX).

Consider a VAR model and let define, n and n_e respectively the number of endogenous exogenous variables:

$$y_t = c + B_1 y_{t-1} + \cdots + B_p y_{t-p} + E_1 x_t + \cdots + E_q x_{t-q} + u_t \; \forall t \in [1, \cdots, T] \tag{2.1}$$

where Y_t is an $n \times 1$ vector of observed endogenous variables, X_t is an $n_e \times 1$ vector of observed exogenous variables and c is an $n \times 1$ vectors of constants; $B_i, i = 1, \cdots, p$ are $n \times n$ matrices of coefficients; $E_i, i = 1, \cdots, q$ are $n_e \times n_e$ matrices of coefficients for the exogenous variables; u_t are heteroskedastic shocks associated with the VAR equations.
In particular, we assume that the covariance matrix of u_t is defined as:

$$cov\left(u_t\right) = \Sigma_t^{-1} = A^{-1} H_t A^{-1'} \tag{2.2}$$

where $H_t = diag\left(\sigma_{1,t}^2 \frac{1}{\lambda_{1,t}}, \sigma_{2,t}^2 \frac{1}{\lambda_{2,t}}, \cdots, \sigma_{n,t}^2 \frac{1}{\lambda_{n,t}}\right)$ with $\sigma_{k,t} = \ln \sigma_{k,t-1} + s_{k,t}$ for $k = 1, \cdots, n$, where $s_{k,t}$ is the error terms associated with the volatility, and $var(s_{k,t}) = g_k$, where g_k is the the shocks to the volatility transition from Cogley and Sargent (2005).

As shown by Geweke (1993), assuming a Gamma prior for $\lambda_{k,t}$ of the form $p\left(\lambda_k\right) = \prod_{t=1}^{T} p\left(\lambda_{k,t}\right) = \prod_{t=1}^{T} \widetilde{\Gamma}\left(1, v_{\lambda,k}\right)$ leads to a scale mixture of normals for the orthogonal residuals $\widetilde{\epsilon_t} = Au_t$ where $\widetilde{\epsilon_t} = \{\widetilde{\epsilon_{1,t}}, \widetilde{\epsilon_{2,t}}, \cdots, \widetilde{\epsilon_{n,t}}\}$ and $cov\left(\widetilde{\epsilon_t}\right) = H_t$ Note that $\widetilde{\Gamma}\left(a, b\right)$ denotes a Gamma density with mean a and b degrees of freedom b.

As explained in Primiceri (2005) and Chiu et al. (2017b), the stochastic volatility allows capturing possible heteroscedasticity of the shocks and potential nonlinearities in the dynamic relationships of the model variables, which are related to the low-frequency variations in the volatility. Baumeister et al. (forthcoming) also rely on such extensions of the linear specification. In order to deal with high-frequency variations in volatility, one way is to consider Student's t-distribution in the shock structure. As these variations are often of extreme magnitudes, this specification can deal efficiently with outliers and extreme events.[16] Therefore, the tVARVOLX model allows us to capture both transient and persistent shifts in volatility.[17]

Furthermore, in order to assess whether persistent or transient shifts in volatility most characterizes the data, we consider the two following restricted models: the VARVOLX model allowing for stochastic volatility only and the tVARX model allowing for student's errors only. The BVARX model corresponds therefore to the most restricted model when fat tails and stochastic volatility are not allowed.

2.3.2 Variable selection

As each model presents a large number of parameters to estimate, a variable selection procedure can be useful to reduce overfitting risk in our out-of-sample forecasting exercise. We have implemented variable selection for all our specifications. As in Korobilis (2013) it is possible to choose from the following three types of priors: Ridge regression prior, Minnesota (Litterman) prior and Hierarchical Bayes Shrinkage prior for the penalization of coefficients. In order to ensure compatibility of the results with the models without variable selection, we retain Minnesota (Litterman) prior. To summarize the following specifications have been implemented using variable selection as in Korobilis (2013): BVARX, tVARX, VARVOLX and tVARVOLX. The pseudo-algorithm for variable selection is available in the Appendix.

2.4 Estimation algorithms and forecasting procedure

We estimated all our models using Markov Chain Monte Carlo (MCMC) and Gibbs sampler. The VAR parameter priors are calibrated as in Bańbura et al. (2010)[18].

[16] Jacquier et al. (2004) provide a detailed analysis of this issue in a univariate framework.

[17] For more details about tVARVOL model see Chiu et al. (2017b).

[18] More details about prior and hyperparameter choice for different VAR specifications (tVAR, VARVOL and tVARVOL) can found in Chiu et al. (2017b)

Our reported results are based on 20,000 Gibbs replications discarding the first 15000 as burn-in. Figure 2.3 in Appendix shows a little fluctuation in recursive mean of some selected parameters, indicating the convergence of the Gibbs algorithm for the retained draws.

We proceed by assessing the forecasting performance of our models considered above by producing a pseudo out-of-sample forecasts[19]. The four models are estimated recursively from August 2013 to August 2018. At each iteration, we construct the forecast density for the models.

$$P\left(\hat{y}_{t+h} \mid y_t\right) = \int P\left(\hat{y}_{t+h} \mid y_t, \Psi_{t+h}\right) P\left(\hat{\Psi}_{t+h} \mid \Psi_t, y_t\right) P\left(\hat{\Psi}_t \mid y_t\right) \mathrm{d}\Psi \tag{2.3}$$

where $h = 1, 2, \cdots, 12$ and Ψ denotes the model parameters. $P\left(\hat{\Psi}_t \mid y_t\right)$ represents the posterior density of the parameters which is obtained via the MCMC simulation. $P\left(\hat{y}_{t+h} \mid y_t, \Psi_{t+h}\right)$ and $P\left(\hat{\Psi}_{t+h} \mid \Psi_t, y_t\right)$ denote the density forecast of the data and the parameters that can be obtained by simulation. The point forecast is therefore the mean of the forecast density. The density forecasts, which are evaluated using log scores (LS), are defined as follows:

$$LS_t = \ln P\left(y_{t+h}\right) \tag{2.4}$$

where $P\left(y_{t+h}\right)$ denotes the forecast density evaluated at data observations. A higher value for LS_t suggests a more accurate density forecast. As in Chiu et al. (2017b) we employ kernel methods to estimate the density and distribution function of the forecasts. This addresses potential non-linearity in the forecast distribution.

Considering the accuracy of point forecasts, we use the *Mean Square Prediction Errors* (*MSPE*s). The *MSPE* for $h = 1, \cdots, 12$ is computed as:

$$MSPE_h = \frac{1}{T-R} \sum_{t=R}^{T-1} \left(\hat{y}_{h,t+1|t} - y_{h,t+1}\right)^2 \tag{2.5}$$

[19] All the estimations and computations presented in this paper were carried out using Julia. As recursive estimation of different models can be very time consuming, we take advantage of the computing clusters through the use of parallel computing in Julia.

where T is the number of observations, R is the length of the rolling window and $\hat{y}_{h,t+1|t}$ are the individual forecast of monthly average of real Henry Hub natural gas price forecasts and $y_{h,t+1}$ are monthly average of real Henry Hub natural gas price forecasts. For this purpose we report the *MSPE* of each specification as well as the *MSPE* ratio relative to no change forecasts.

Moreover, in order to statistically assess the best forecasting model, we perform both conditional and unconditional pairwise model comparisons based on forecasting performance test procedure of Giacomini and White (2006). This procedure is particularly relevant for both point and density forecasts when using recursive forecasting on real-time data.

2.5 Empirical results

In this section, we first show the relative forecasting performance when using real-time data relative to post-revised data. Second, we highlight the importance of including the temperature within BVAR frameworks in giving more accurate forecasts of real gas price. We also compare models that represent different specifications taking into account fat-tails and stochastic volatility in term of both density and point forecasting.

2.5.1 Do Real-Time data improve forecast accuracy?

In order to assess the usefulness of using real-time data, we do perform the same forecasting exercise using post-revised data. To do this, we measure ratios of MSPE of model forecasts using real-time data relative to model forecasts using post-revised data.

Table 2.1 summarizes the real gas price forecast accuracy of both univariate and multivariate Bayesian autoregressive models. BVAR models include market drivers of real gas price, namely the U.S. Natural Gas production (P) for the supply side and the U.S. Natural gas consumption (C) and an estimate of the U.S. Natural Gas Inventories (S) demand side, the real West Texas Intermediate oil prices (WTI). As is common since Baumeister and Kilian

(2012), we consider the estimated recursive MSPE relative to that of the no-change forecast for one-month, three-month, six-month, nine-month and twelve-month ahead forecasts.

Table 2.1: Recursive MSPE of forecasting accuracy

	Panel A: Real-Time				Panel B: Real-time over Post-Revised			
	BAR	BVAR			BAR	BVAR		
		P,C,HH,S	WTI,P,HH,C	WTI,P,C,HH,S		P,C,HH,S	WTI,P,HH,C	WTI,P,C,HH,S
1M	1.1934	1.1265	1.0530	1.0307	1.0000	1.1412	1.0621	1.0341
3M	1.3037	1.0029	0.9289**	0.9121*	1.0000	0.9602*	0.9410*	0.9466*
6M	1.4487	0.8563**	0.7785***	0.7723**	1.0000	0.9049**	0.8044**	0.8274**
9M	1.6103	0.6988 ***	0.6371***	0.6368***	1.0000	0.7591***	0.6736***	0.6732***
12M	1.5902	0.5878***	0.5288***	0.5295***	1.0000	0.6583***	0.5732***	0.6009***

Notes: BAR is estimated using 12 lags. For BAR and BVAR, the best forecast accuracy is obtained using 12 lags for the exogenous variable. Panel A presents ratios of each model's MSPE to the baseline no change forecast successively evaluated for the entire evaluation period (August, 2013 – August, 2018). Panel B presents MSPE ratios of model including real-time data over those including post-revised data. The notation BAR and BVAR, refers respectively to univariate Bayesian autoregression model and Bayesian vector autoregression. Point forecasts are produced with recursive estimation of the models. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The notation C, HH, P, S and WTI refers respectively to US Natural gas consumption, Monthly average of Henry Hub Natural Gas real prices, US Natural Gas production, US Natural Gas Inventories and Monthly average real West Texas Intermediate oil prices. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini-White (2006) test.

Not surprisingly and as shown in Baumeister and Kilian (2012) for oil market, results from Table 2.1 Panel B shows models that include real-time data indicate a statistically significant improvement over models that include post-revised data; for all models and for all horizons, except for one month horizon, MSPE ratios are less than one, improvement reaching for instance 40% for 12 month horizon. This substantial differences in forecasts support our hypothesis that using real-time data solve at least partly the forecasting problems caused by the existence of data revisions.

Moreover, results from Table 2.1 (Panel A and B) provide a strong evidence that BVAR models give higher forecast accuracy than univariate model, confirming our choice of multivariate Bayesian framework. Apart from results at one-month horizon, the forecast accuracy gains using BVAR models with respect to no-change are between 7% and 47% for the remaining horizons. Accuracy gains are more pronounced when comparing these model forecasts with BAR model. For example, the BVAR model reduces the MSPE by between 7% and 100% at all horizons relative to BAR model.

Interestingly, a comparison across BVAR models including either real-time or revised data reveals the benefit of including WTI in term of forecast accuracy. Including WTI reduce relative

MSPE by 21% in average for all horizons, whereas, including gas inventories reduce relative MSPE by 14% only for all horizons. Moreover, the five-variable BVAR model including both WTI and gas inventories gives roughly the same forecast performance as the four-variable BVAR including WTI only. Accordingly, and for reasons which will be explained in the next section, our benchmark model will include WTI and exclude U.S. gas inventories.[20]

To summarize, linear Bayesian VAR including variables that are quite similar to Baumeister and Kilian (2012) deliver highly competitive forecasts relative to the no-change prediction. This is an important result that confirms the prominence of our new data set and the fact that fundamental variables for the natural gas market are, collectively taken, able to generate good forecasts of the future gas price.

2.5.2 The role of temperature

We now wish to take advantage of the strong relationship between temperature and gas consumption. Do models including temperature provide more accurate real gas price forecasts? This hypothesis is tested by including the temperature into the vector autoregressive model as an exogenous variable in the form of a BVARX.

Linear Gaussian Bayesian VAR

In Table 2.2 we report the performance of BVARXs with the Heating Degree Days (HDD) and with our measure of temperature, the normalized cumulated heating degree days (CHDD), over our benchmark model (BVAR). A remarkable finding is that BVARX models which incorporate CHDD evidently dominate in term of MSPE. The gains generated by BVARX model are significant with a magnitude between 5% and 15% for all horizons relative to the BVAR model. We also find that this forecasting accuracy beneficial prevails even when using post-revised data as the BVARX outperforms the benchmark model with an accuracy gain between

[20]The use of Cholesky decomposition implies the choice of variable order. WTI is ordered first in the VAR in light of the findings in Kilian and Vega (2011). Then, for robustness, we estimate different models with different variable orders and results remain unchanged.

3% and 12% (Panel B in Table 2.2).

Importantly and as expected, integrating Heating Degree Days (HDD) leads instead to a much worse forecasting ability of BVAR model, suggesting that the noise added from using this variable does more damage than the good that comes from the strong correlation that it has with gas consumption (60%). This confirms our hypothesis that HDD contains information which is already comprised in gas consumption and only extraordinary weather conditions should explain real gas price formation.

Moreover, results from relative MSPE are further re-enforced by the density forecast accuracy test. Table 2.9, in Appendix D, gives results of the log score of BVARX including our measure of temperature relative to the BVAR model, so that positive values indicate an improvement over our benchmark. Results from Table 2.2 show that the BVARX models outperform the BVAR model leading to a generally significant forecast improvement by between 5% and 15% when real-time data are used. However, the magnitude of forecast improvement of BVARX based on revised data is smaller, between 3% and 12%, and non-significant.

Our results confirm that the inclusion of temperature data is fruitful at all horizons thereby confirming the important role of considering the demand side for modeling gas prices.

Table 2.2: Recursive MSPE ratio relative to BVAR as the benchmark

	Panel A: Real-Time				Panel B: Post-Revised			
	BVAR	BARX	BVARX		BVAR	BARX	BVARX	
X		CHDD	HDD	CHDD		CHDD	HDD	CHDD
1M	0.0756	3.4404	1.5058	0.9516*	0.0711	3.6542	1.6155	0.9682*
3M	0.2663	2.5093	1.3310	0.8997*	0.2833	2.3613	1.5026	0.9495*
6M	0.3985	1.8531	1.4328	0.8479**	0.4954	1.4908	1.1487	0.9266*
9M	0.4105	2.5401	1.4875	0.8711**	0.6093	1.7110	1.1252	0.9028*
12M	0.3878	2.0271	1.6064	0.9271*	0.6765	1.1620	1.1242	0.8808*

Notes: BARX is estimated using 12 lags. Panel A (resp. B) presents MSPE ratios of model including real-time data (resp. including post-revised data). For BVAR only MSPE is provided. We perform recursive forecasts with a first estimation on a sample covering data up to July 2013 and subsequent estimations adding one more observation each time until August, 2018. The notation BARX, BVAR and BVARX, refers respectively to univariate Bayesian autoregression model with exogenous variable, Bayesian vector autoregression without and with exogenous variable. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The multivariate endogenous dataset which is use to calibrate is; US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average ofHenry Hub Natural Gas real prices, US Natural gas consumption. The notation HDD and CHDD refers respectively to Heating Degree Days and Cumulative Heating Degree Days. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini-White (2006) test.

Considering stochastic volatility and fat tails

In this section we retain the BVARX with our measure of temperature as exogenous variable as our benchmark model. This is motivated by two reasons. First, because this model provides the largest forecast performance with respect to BVAR models. Second, This allows us to evaluate the usefulness of our measure on temperature when adjusting the base linear and Gaussian model to account for stochastic volatility and fat tails. In addition, we implement the variable selection method, explained in section 3, with the objective of further improving forecasting performance of different models.

Before evaluating the forecast performance of different models, we begin by comparing their fitting specification (adequacy) using the marginal likelihood.

Table 2.3: Log-likelihood

	BVARX		tVARX		VARVOLX		tVARVOLX	
VS	No	Yes	No	Yes	No	Yes	No	Yes
Log-likelihood	-1988	-1934	-1461	-1458	-1042	-1040	-1039	-1034

The log marginal likelihood values are estimated for each model using the full sample with 20.000 iterations. The BVARX has the lowest log marginal likelihood.

Table 2.3 gives the estimated values of the log marginal likelihood for each model using the full sample.[21] Results show that the combination of fat-tail and stochastic volatility within the tVARVOLX model has the highest estimated marginal likelihood delivering therefore the best in-sample fitting specification. Even though allowing for fat-tail improves the fit relative to the BVARX model, allowing for stochastic volatility gives much more improvement. This may indicate that stochastic volatility is more adequate for our data.

As regards the forecasting performance, Table 2.4 reports recursive MSPE of alternative models relative to the benchmark model with and without variable selection. Note that, elements reported in the first column correspond to MSPE measures of the benchmark model (BVARX) without variable selection.

Three points emerge from Table 2.4. First, allowing for fat tails appears to be beneficial as tVARX model offers an MSPE reduction as large as 14% at horizon 12 without variable

[21] Estimation method of the log marginal likelihood is explained in Chiu et al. (2017b).

Table 2.4: Recursive MSPE ratio relative to BVARX

	BVARX		tVARX		VARVOLX		tVARVOLX	
VS	No	Yes	No	Yes	No	Yes	No	Yes
1M	0.0720	0.9931*	0.9929	0.9600*	0.9625*	0.9699*	0.9685*	0.9631*
3M	0.2396	0.9766*	0.9395*	0.9191*	0.8797**	0.8800**	0.8754**	0.8725**
6M	0.3379	0.9612*	0.9213*	0.9209*	0.8735**	0.8734**	0.8659**	0.8630**
9M	0.3576	0.9521*	0.8922*	0.9141*	0.8632**	0.8630**	0.8524**	0.8496**
12M	0.3596	0.9487*	0.8588**	0.9034*	0.8511**	0.8509**	0.8374**	0.8344**

Notes: This table presents ratios of each model's MSPE to the baseline BVARX models successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The forecasts are produced with recursive estimation of the models. For BVARX only MSPE is provided. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The notation BVARX, VARVOLX, tVARX and tVARVOLX refer respectively to Bayesian vector autoregression with exogenous variable, Bayesian vector autoregression with exogenous variable and stochastic volatility, Bayesian vector autoregression with exogenous variable and fat tails (t-distributed shock), Bayesian vector autoregression with exogenous variable, stochastic volatility and fat tails (t-distributed shock). Yes (resp. No) in line VS (Variable selection) means (resp. No) Bayesian shrinkage in the VAR coefficient. The multivariate endogenous dataset which is use to calibrate is; Real-Time US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, Real-Time US Natural gas consumption with Cumulative Heating Degree Days is used as exogenous regressors. Only 12 lags BVARX models is reported because is the best lag parameter value.Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate RMSE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test.

selection (only 10% when variable selection is applied), about 8% for horizons 3, 6 and 9, and between 0 and 4% at horizon 1 for model without and with variable selection, respectively. These improvements nevertheless remain borderline statistically significant. It appears that variable selection is only beneficial at short horizons 1 and 2. A second notable point is that modeling low-frequency movements in volatility appears to be useful at all horizons, especially at longer horizons. Compared to tVARX model, VARVOLX makes larger and more significant improvements. This leads to think that gas market variables are more characterized by persistent than transient shifts in volatility. Third, an interesting observation in Table 2.4 is that the combination of fat tails and stochastic volatility (tVARVOLX) does deliver the best performance among all models considered. The TVARVOLX dominates our benchmark significantly with MSPE reduction between 13% and 16% at horizons 3, 6, 9 and 12. As with tVARX and VARVOLX, marginally larger MSPE reduction is observed at horizon 1.

Further robustness analysis is included in Table 2.10 (Appendix D), which generally corroborates our previous findings. Log score comparisons are in line with the relative MSPE measures confirming that BVARX forecast is strictly dominated when fat-tails and stochastic volatility are taken into account. Improvements are highly significant at longer horizons, 6, 9 and 12, but not significant at horizon 3. At horizons 1, however, only tVARVOLX without variable selection hardly succeeds to provide a marginal improvement. Once again, VARVOLX produces better forecasting performance than tVAR model, both in magnitude and significance, and the largest forecast gain is provided by tVARVOLX with improvement reaching for example 58% at horizon 12.

It is not implausible that the forecasting performance of BVARs allowing for fat-tails and stochastic volatility is totally due to their own specifications and completely unrelated to our measure of temperature. To test this hypothesis, we compare forecasting performance of these models with their corresponding models without temperature.

Table 2.5: Recursive MPSE ratio relative to each specifications without exogenous variable

	BVARX /BVAR		TVARX /TVAR		VARVOLX /VARVOL		TVARVOLX /TVARVOL	
VS	No	Yes	No	Yes	No	Yes	No	Yes
1M	0.9516*	1.0189	1.003	0.9960	0.9941	0.9623	1.0146	1.0210
3M	0.8997*	0.8901*	0.8999*	0.8980*	0.9021*	0.8956*	0.9481*	0.9473*
6M	0.8479**	0.8657**	0.9997	0.9992	0.9999	0.9968	0.9247*	0.9209*
9M	0.8711**	0.8459**	0.8634**	0.8592**	0.9167*	0.8955*	0.8944*	0.8971*
12M	0.9271*	0.9303*	0.9547*	0.9293*	0.8960*	0.9001*	0.8623**	0.8664**

Notes: This table presents ratios of each model's MSPE to the associated model without exogenous variable models successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The forecasts are produced with recursive estimation of the models. The notation BVAR (X), VARVOL(X), tVAR(X) and tVARVOL(X) refer respectively to Bayesian vector autoregression (with exogenous variable), Bayesian vector autoregression (with exogenous variable) and stochastic volatility, Bayesian vector autoregression (with exogenous variable) and fat tails (t-distributed shock), Bayesian vector autoregression (with exogenous variable), stochastic volatility and fat tails (t-distributed shock). Yes (resp. No) in line VS (Variable selection) means (resp. No) Bayesian shrinkage in the VAR coefficient. The multivariate endogenous dataset which is use to calibrate is; Real-Time US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, Real-Time US Natural gas consumption with Cumulative Heating Degree Days is used as exogenous regressors. Only 12 lags BVARX models is reported because is the best lag parameter value. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate RMSE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test.

Relative MSPE measures from Table 2.5 bring out the relevance of our measure of temperature as all models including CHDD dominate their corresponding model without the exogenous variable, with one exception for tVARX and VARVOLX at horizon 6 and anothor for TVARVOLX at horizon 1. Log score measures given in Table 2.11 show that even at these horizons, models incorporating CHDD remain the best performing competitors, though forecasting gains are not statistically significant.

As pointed out by Korobilis (2013), the higher the number of variables in the model, the more useful penalization within variable selection would be. In our study, variable selection has in many cases improved the performance of forecasts, despite the small number of variables included in the VAR models. We estimate our models with different values of the hyper-parameter (π_{0j}) that parameterizes the proportion of the predictors that should be in the final model. We select thus values between 0.5 (non-informative), meaning that half of the BVAR coefficients should be restricted, and 0.8 when only 20% of coefficients will be restricted. We finally retained a value of 0.8 for the reported result estimations as it gives the best forecasting gain. This is also consistent with Korobilis (2013)'s assertion that, for small BVAR, the choice of non-informative prior for this hyper-parameter leads probably to too many restricted BVAR coefficients, worsening therefore the model forecast performance.

2.6 Concluding remarks

The U.S. natural gas price and its evolution is of utmost importance for the U.S. economy and decent forecasts of the real price of gas are useful for market participants as well as for consumers, regulators and central bankers.

In this paper, we estimate various Bayesian VAR models following the seminal contribution by Baumeister and Kilian (2012) and more recently Baumeister et al. (forthcoming) with an emphasize on the demand side which is shown to be highly relevant in the case of gas. Moreover, the use of temperature data that is made possible by the regional feature of the U.S. gas market permits to further include a true real-time proxy for demand in our econometric specifications.

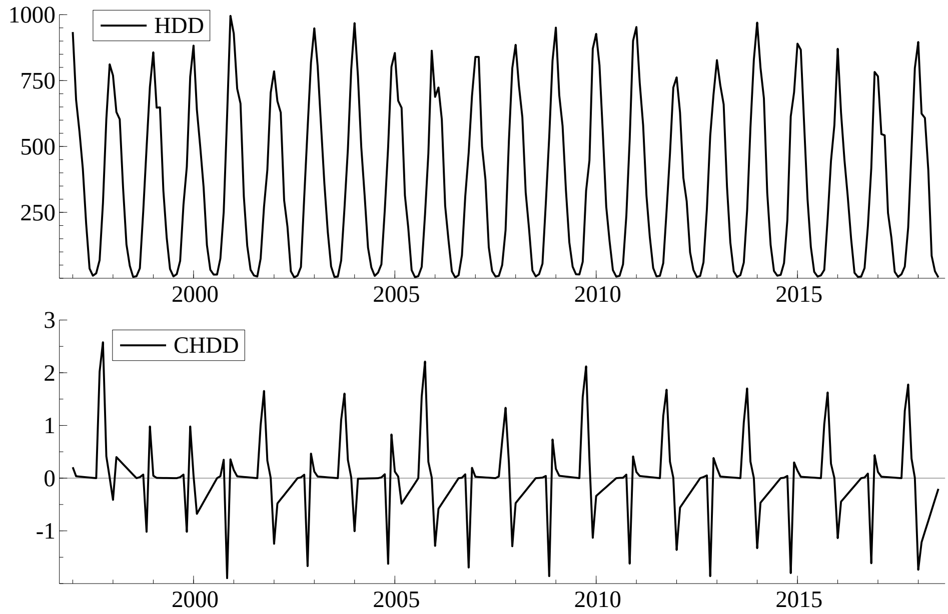
Our results, that are the first of this kind for natural gas and are likely to serve as future benchmarks for economic research in the field of energy economics, point to large improvements relative to the no-change forecast. We provide evidence that models that comprise temperature and allow for stochastic volatility and fat tails deliver the best forecasts. The role of temperature, in particular, is a novel finding that deserves to be further investigated.

Another strand of the literature initiated in Alquist and Kilian (2010b) deals with the information content of futures prices to make predictions about commodity prices. While we do not elaborate on this possibility for natural gas in the present paper, a related approach in Thomas (2020) is noteworthy. More specifically, Thomas (2020) relies on a non causal bivariate VAR for crude oil and natural gas. In his model, the two variables under consideration are the energy price and the convenience yield. The latter variable is included on the economic ground that it does proxy for expectations in the derivatives market. Surprisingly, such a very parsimonious model exhibits highly accurate forecasts at horizons up to several months.

2.7 Appendix

2.7.1 Data

Figure 2.2: Our measure of temperature



2.7.2 Appendix B: Empirical convergence results

The various figures in 2.3 represents the recursive mean of the variation in the coefficients of the autoregressive part of the Bayesian model on the iterations retained at the end of the Gibbs sample (5000 iterations). According to these figures, the variations in the value of the estimate of each coefficient are minimal, which ensures the convergence of the model.

Figure 2.3: Mean recursive estimations of VAR coefficients for different specification

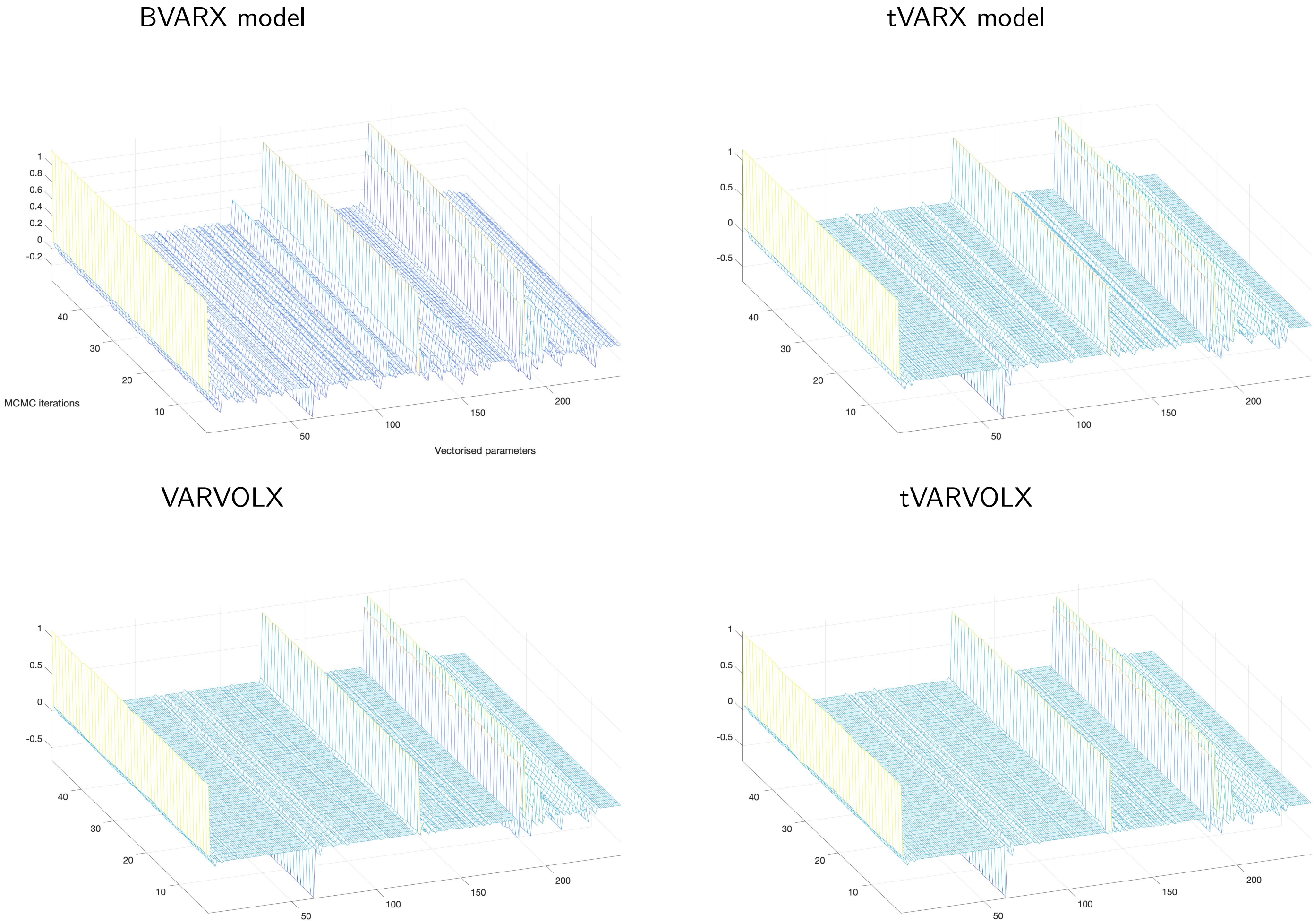
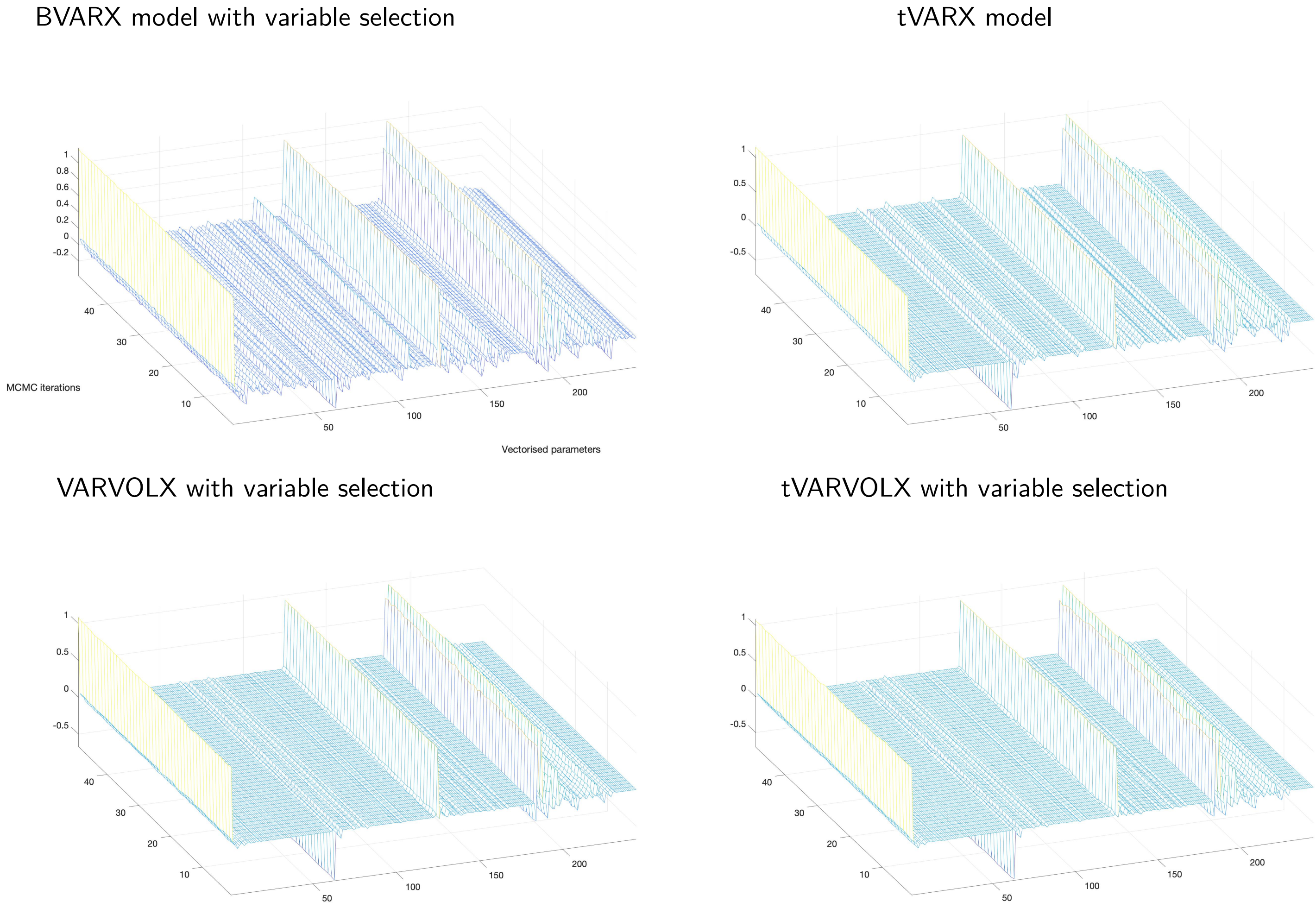


Figure 2.4: Mean recursive estimations of VAR coefficients for different specification



2.7.3 Appendix C: Temperature and post-revised data

Table 2.6: Recursive MSPE ratio relative to the no-change forecast with post-revised data

	BAR	BVAR		
		P,C,HH,S	WTI,P,HH,C	WTI,P,C,HH,S
1M	1.1934	0.9871	0.9914	0.9967
3M	1.3037	1.0443	0.9871	0.9635
6M	1.448	0.9462	0.9678	0.9334*
9M	1.6103	0.9205*	0.9458	0.9459
12M	1.5902	0.8929**	0.9225*	0.8811**

Notes: BAR is estimated using 12 lags. The table presents ratios of each model's MSPE to the baseline no change forecast successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The notation BAR and BVAR, refers to univariate Bayesian autoregression model and Bayesian vector autoregression. Point forecasts are produced with recursive estimation of the models. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The notation C, HH, P, S and WTI refers respectively to US Natural gas consumption, Monthly average of Henry Hub Natural Gas real prices, US Natural Gas production, US Natural Gas Inventories and Monthly average real West Texas Intermediate oil prices. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini-White (2006) test.

Table 2.7: Recursive MSPE ratio relative to the no-change forecast (Temperature)

	Panel A: Real-Time				Panel B: Post-Revised			
	BVAR	BARX	BVARX		BVAR	BARX	BVARX	
X		CHDD	HDD	CHDD		CHDD	HDD	CHDD
1M	1.0530	3.6228	1.5857	1.0021	0.9914	3.6228	1.6017	0.9606*
3M	0.9289*	2.3309	1.3310	0.8339**	0.9871**	2.3309	1.4833	0.9373*
6M	0.7785***	1.4427	1.0898	0.6601***	0.9678*	1.4427	1.1041	0.8967*
9M	0.6371***	1.6183	0.9477*	0.5550***	0.9458*	1.6183	1.0643	0.8539**
12M	0.5288***	1.0720	0.8495**	0.4903***	0.9225*	1.0720	1.0371	0.8124**

Notes: BARX is estimated using 12 lags. For BARX and BVARX, the best forecast accuracy is obtained using 12 lags for the exogenous variable. Panel A (resp. Panel B) presents ratios of each model's MSPE including real-time data (resp. including post-revised data) to the baseline no change model forecast successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The notation BARX, BVAR and BVARX, refers respectively to univariate Bayesian autoregression model with exogenous variable, Bayesian vector autoregression without and with exogenous variable. Point forecasts are produced with recursive estimation of the models. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The multivariate endogenous dataset which is use to calibrate is; US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, US Natural gas consumption. The notation HDD and CHDD refers respectively to Heating Degree Days and Cumulative Heating Degree Days. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini-White (2006) test.

2.7.4 Appendix D: Density forecast accuracy

Table 2.8: Recursive average percentage improvement in Log-Score

Real-time over Post-Revised			
	BVAR		
	P,C,HH,S	WTI,P,HH,C	WTI,P,C,HH,S
1M	-3.62	-5.75	-5.39
3M	2.85	2.88	6.54
6M	4.77	1.49	1.03
9M	7.29*	5.52	1.41
12M	7.46*	18.00**	10.30

Notes: The table presents the percentage values of log predictive density scores of model including real-time data model relatively to those including post-revised data successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The notation BVAR refers to univariate Bayesian vector autoregression. Density forecasts are produced with recursive estimation of the models. Positive entries indicate that Log-Score from the indicated model are improved (in %) than Log-Score from the associated baseline model. The multivariate endogenous dataset which is use to calibrate is; US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, US Natural gas consumption. The notation HDD and CHDD refers respectively to Heating Degree Days and Cumulative Heating Degree Days. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate Log-Score variations are significantly different from 0 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini-White (2006) test.

Table 2.9: Recursive average percentage improvement in Log-Score relative to BVAR as the benchmark

	Panel A: Real-Time				Panel B: Post-Revised			
	BVAR	BARX	BVARX		BVAR	BARX	BVARX	
X		CHDD	HDD	CHDD		CHDD	HDD	CHDD
1M	0.7062	-86.024	-24.15	2.10	0.7417	-82.78	-27.98	-0.01
3M	0.5930	-765.26	-11.49	5.77	0.5186	-731.26	-12.07	4.20
6M	0.4399	-2826.83	-12.98	10.38*	0.4386	-2857.98	-10.43	9.99*
9M	0.3707	$-\infty$	-17.34	12.67**	0.3516	$-\infty$	-15.95	8.64*
12M	0.3314	$-\infty$	-28.66	8.20*	0.2806	$-\infty$	-20.05	8.35*

Notes: BARX is estimated using 12 lags. Panel A (resp. Panel B) presents the percentage values of log predictive density scores of model including real-time data (resp. including post-revised data) model relatively to BVAR. For BVAR only Log-Score is provided. We perform recursive forecasts with a first estimation on a sample covering data up to Jully 2013 and subsequent estimations adding one more observation each time until August, 2018. The notation BARX, BVAR and BVARX, refers respectively to univariate Bayesian autoregression model with exogenous variable, Bayesian vector autoregression without and with exogenous variable. Positive entries indicate that Log-Score from the indicated model are improved (in %) than Log-Score from the associated baseline model. The multivariate endogenous dataset which is use to calibrate is; US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, US Natural gas consumption. The notation HDD and CHDD refers respectively to Heating Degree Days and Cumulative Heating Degree Days. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate Log-score variations are significantly different from 0 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini-White (2006) test.

Table 2.10: Recursive average percentage improvement in Log-Score relative to BVARX as the benchmark

	BVARX		tVARX		VARVOLX		tVARVOLX	
VS	No	Yes	No	Yes	No	Yes	No	Yes
1M	0.7210	1.22	-5.82	-5.78	-2.52	-3.16	1.4	-4.16
3M	0.6272	5.53	2.98	2.79	1.23	3.89	3.40	5.02
6M	0.4956	2.84	11.03^{*}	9.50^{*}	18.24^{***}	21.11^{***}	24.28^{***}	23.87^{***}
9M	0.4092	2.10	19.09^{***}	18.65^{***}	32.16^{***}	33.28^{***}	37.00^{***}	40.42^{***}
12M	0.3610	6.00	30.68^{***}	31.20^{***}	48.33^{***}	48.85^{***}	53.79^{***}	58.22^{***}

Notes: The table reports the percentage values of log predictive density scores model relatively to the baseline models. The baseline model is BVARX model without variable selection. For BVARX only predictive density scores is provided. Positive entries indicate that Log-Score from the indicated model are improved (in %) than Log-Score from the associated baseline model The forecasts are produced with recursive estimation of the models successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The notation BVARX, VARVOLX, tVARX and tVARVOLX refer respectively to Bayesian vector autoregression with exogenous variable, Bayesian vector autoregression with exogenous variable and stochastic volatility, Bayesian vector autoregression with exogenous variable and fat tails (t-distributed shock), Bayesian vector autoregression with exogenous variable, stochastic volatility and fat tails (t-distributed shock). Yes (resp. No) in line VS (Variable selection) means (resp. No) Bayesian shrinkage in the VAR coefficient. The multivariate endogenous dataset which is use to calibrate is; Real-Time US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, Real-Time US Natural gas consumption with Cumulative Heating Degree Days is used as exogenous regressors. Only 12 lags BVARX models is reported because is the best lag parameter Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate Log-score variations are significantly different from 0 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test.

Table 2.11: Recursive average percentage improvement in Log-Score relative to each specifications without exogenous variable

	BVARX /BVAR		TVARX /TVAR		VARVOLX /VARVOL		TVARVOLX /TVARVOL	
VS	No	Yes	No	Yes	No	Yes	No	Yes
1M	2.10	2.89	0.43	-0.38	0.39	1.00	1.70	2.60
3M	5.77^{*}	7.47^{*}	1.19	1.00	2.29	0.47	6.61^{*}	4.82
6M	12.67^{**}	11.93^{**}	0.13	1.85	3.61	4.00	17.44^{**}	17.01^{**}
9M	10.38^{**}	14.33^{**}	0.53	1.08	2.25	5.50	7.98^{*}	18.09^{**}
12M	8.93^{*}	8.07^{*}	2.06	1.29	4.08	2.95	24.15^{***}	21.78^{***}

Notes: The table reports the percentage values of log predictive density scores model relatively to the associated model without exogenous variable models successively evaluated for the entire evaluation period (August, 2013 – August, 2018). The forecasts are produced with recursive estimation of the models. Positive entries indicate that Log-Score from the indicated model are improved (in %) than Log-Score from the associated baseline model. The notation BVAR (X), VARVOL(X), tVAR(X) and tVARVOL(X) refer respectively to Bayesian vector autoregression (with exogenous variable), Bayesian vector autoregression (with exogenous variable) and stochastic volatility, Bayesian vector autoregression (with exogenous variable) and fat tails (t-distributed shock), Bayesian vector autoregression (with exogenous variable), stochastic volatility and fat tails (t-distributed shock). Yes (resp. No) in line VS (Variable selection) means (resp. No) Bayesian shrinkage in the VAR coefficient. The multivariate endogenous dataset which is use to calibrate is; Real-Time US Natural Gas production, Monthly average real prices of West Texas Intermediate, Monthly average of Henry Hub Natural Gas real prices, Real-Time US Natural gas consumption with Cumulative Heating Degree Days is used as exogenous regressors. Only 12 lags BVARX models is reported because is the best lag parameter value. Each model is estimated using 11,000 iterations, with 6,000 burns.***, ** and * indicate Log-score variations are significantly different from 0 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test.

Chapter 3

A Structural Non-causal VAR Model of the Global Oil Market: the Role of Oil Supply News Shocks

[This chapter is based on Moussa and Thomas (2020)]

3.1 Introduction

The global oil market has experienced unprecedented turbulence due to the collapse in demand caused by the coronavirus pandemic, despite the OPEC's reaction. Understanding how demand and supply shocks affect the (real) oil price and how real economy responds to a change in oil price become increasingly prominent. A large number of recent empirical papers that aim to identify and measure the effect of oil supply and demand shocks have employed structural vector autoregressive (SVAR) models. Within SVAR models, different identification strategies have been employed based on either timing (Kilian, 2008a,b, 2009) or sign restrictions (Kilian and Murphy, 2012, 2014b; Lippi and Nobili, 2012; Baumeister and Peersman, 2013; Baumeister and Hamilton, 2019a). Caldara et al. (2019), Zhou (2020) and Känzig (2019), instead, adopt an alternative approach that uses either proxy or external instrumental variable procedures within SVAR model (*proxy* SVAR). These empirical studies (an exhaustive review on this literature is in Kilian and Zhou (2020)) show that oil supply and demand shocks are hard to identify in practice, leading to different contradictory results regarding the predominance of supply and demand shocks in explaining oil price fluctuations.

Even though SVAR models are widely used in most empirical studies on oil market, they may suffer from nonfundamentalness problem which arises when economic agents' information set is larger than the econometrician's one (see Alessi et al. (2011) for a review). When economic agents have expectations of future oil supply or demand, they can use additional information to form such expectations, while the econometrician estimating a VAR model exploits only a limited amount of information not including agent's expectations. As a result of the informational deficiency the reduced form innovations do not contain enough information to recover the structural shocks of interest. This is evident in the context of global oil market context. As with any asset price, oil price would incorporate all the information available to the agent that affect its expectations about future oil supply and demand. Expectations about future oil demand may depend, for example, on one's view of the impact of the coronavirus pandemic on Chinese economy which could have long-lasting and global consequences, or also of the outcome of the trade war between China and the United States. As for expectations about future oil supply, numerous determinants potentially relevant can be related to the political stability in oil producer countries, especially Gulf countries and Iran, to beliefs about an aggressive OPEC price strategy, to large oil discoveries (Arezki et al., 2017) or also to information held by the public about the production capacity constraints in oil producing countries. Therefore, and as argue in Baumeister and Kilian (2016), forward-looking expectations thus imply that it is not the realized demand and supply that have an impact on the oil price, but their levels relative to what it was expected.

Approaches exist to address the nonfundamentalness problem, though each has limitations. The first type of approach is to include additional variables in the VAR system, thus enlarging the econometrician's information set. By including the global inventory as forward-looking variable, Kilian and Murphy (2014b), suggest that expectation shifts in the oil market should be associated with changes in oil inventory. Juvenal and Petrella (2015) use this insight and exploit a larger information set minimising the information deficiency of the SVAR model. These approaches possesses great flexibility without having to model expectations, but do not permit to identify what elements are exactly captured by inventory changes. These later could be associated with speculative demand, expectations about future oil supply shocks, or uncertainty across a number of factors- currencies, geopolitical tensions, changes in the Strategic Petroleum Reserve (Zhou, 2020; Känzig, 2019).

Relatedly, an other solution consists in employing narrative measures and proxies for expectations leading to a more credible identification scheme. Zhou (2020) and Antolín-Díaz and Rubio-Ramírez (2018) employs a narrative sign-restrictions built on Kilian and Murphy (2012) and Kilian and Murphy (2014b)'s reasoning based on key historical events when oil production substantially decreased. Their findings are in line with Kilian and Murphy (2014b) that demand factors are predominant in explaining oil market fluctuations. Caldara et al. (2019) apply the same approach on larger sample of country-specific oil production drops and taking account for more episodes. Their results are more in line with Baumeister and Hamilton (2019a) as they find supply shocks accounting for more relevant part of oil price and quantity fluctuations than in Kilian and Murphy (2014b).

A common feature of all these empirical studies is that they treat nonfundamentalness problem as omitted-variables problem. This response to the nonfundamentalness problem raises the question of the ability of a single variable, namely inventories, factors or proxies to fully catch agents' information set. The second approach is to allow for noninvertible moving average (MA) components. This solution, however, has received little attention and VARs continue to be widely considered. As explained in Chan and Eisenstat (2017), one possible reason is that because estimation and identification of nonfundamental models are rather more complicated than within VARs, many extensions of the basic VAR have developed making it more flexible and therefore more attractive.

The present study aims to complement the existing literature by using Non-Causal VAR (NC-VAR) model (Lanne and Saikkonen, 2013a; Lanne and Luoto, 2016; Gourieroux and

Jasiak, 2017a; Davis and Song, 2020; Nelimarkka, 2017b,a). The aim is to address the above non-fundamentalness and the related identification problems which may arise when only purely autoregressive models are considered. Within NC-VAR model, current observation may depend on both past and future shocks in the system, allowing therefore to capture economic agent expectations of future values of some of the model variables. Allowing VAR model to have both causal and non causal components permits then identifying impulse responses to both structural anticipated and unanticipated shocks.

On the other hand, following the monetary policy literature (Kuttner, 2001; Gürkaynak et al., 2005; Paul, 2020), Känzig (2019) employs an external instrument given by high-frequency price changes in WTI futures around OPEC announcements. This instrument reflects oil supply surprises as it captures the unanticipated part within announcements, and is then used as proxy for the structural oil supply news shocks. There is a strong evidence from this study that oil supply news shocks have significant impact on both oil market and macroeconomic variables, highlighting the role of oil supply expectations. However, even if oil supply surprises are rigorously constructed, the validity of such an instrument may be questioned for the reason that, in contrast to central bank announcements, OPEC announcements about of its production targets are just vague guidelines as there is a systematic deviation from the effective OPEC production. As second concern, which is common to the identifying structural VAR model using high-frequency futures data, is related to the method used to aggregate the daily shock measure from futures data to monthly frequency. Moreover, a doubt regarding the interpretation the OPEC news as oil supply news is pointed out by Kilian and Zhou (2020). According to the authors, based on the variable reactions to the OPEC news shock, this later should rather be interpreted as a shock to price expectations.

In particular, this paper focuses on examining the effects of oil supply news shocks. But instead of including additional proxy measures as in Känzig (2019), oil supply news shocks are identified within the NC-VAR model. We follow the identification technique introduced by Uhlig (2004) and Francis et al. (2014) to measure technology news shocks, and applied by Nelimarkka (2017a) in the context of Non-Causal VAR. Identified oil supply news shocks are then shocks that explain the most of the movements in oil production at a long but finite horizon and have an immediate effect on forward-looking variables. The identified oil supply news shock recover therefore a wider range of information related the oil supply. It could, for instance, be a large oil discoveries, a shock related to the introduction of a new production technology which takes time in order to translate into oil production, and it could also be a shock related to OPEC news announcements.

Using the standard four variables in oil markets structural models, namely global oil production, global economic activity, oil stocks and real oil price and US macroeconomic variables, our analysis yields the following results. First, we find that the Non-Causal VAR is appropriate for modelling oil market variables as nonfundamental representation is supported by the data. Second and relatedly, our identified oil supply news shock is anticipated by forward-looking variables before the shock materialises, highlighting the major role of expectations in propagating the shock. Third, a negative oil supply news shock results in abrupt and permanent fall in global oil production associated with a persistent fall in global economic activity and a permanent increase in oil inventory. However, the reaction of oil price confirms the view that oil supply shocks have a limited effect on oil price. Finally, we show that the news shock about oil supply shortfalls do have macroeconomic consequences as it causes a substantial decline in US industrial production. Moreover, consumer and business confidence indicators decline significantly before the shock materialises, confirming the prominent role of the expectation channel.

The rest of the paper is structured as follows. In the following Section we provide a simple stylised model to illustrate how nonfundamentalness arises out in the presence of lagged shock effects of oil supply on observables. This provides a theoretical motivation for the empirical investigations that follow. Section 3 describes the Non-Causal VAR model and outlines its estimation. Section 4 presents the estimation results. Finally, Section 5 concludes and discusses avenues for future research.

3.2 Stylised model for oil market with news shock

"If oil production was expected to decline, for example, but did not because of a positive oil supply shock, then this shock would trigger an additional [downward] adjustment of the price of oil without a change in observed oil production". Baumeister and Kilian (2016) (p. 135).

This quotation emphasizes well the role of expectations as a major determinant of oil price movements. Indeed, agents receive advanced information or signal about the future oil supply or demand. We refer, therefore the oil supply news as the exogenous changes that alter information set on which the agents base their expectation. In this sens oil supply news shock can be viewed as a technology news shock common in the macroeconomic literature (Beaudry and Portier, 2006; Barsky and Sims, 2011; Forni et al., 2014; Beaudry and Portier, 2014, among others). The news shock recover therefore a wide range of information related the oil supply. It could, for instance, be a large oil discoveries, or a shock related to the introduction of a new production technology which takes time in order to translate into oil production. Alternatively, it could also be a shock related to OPEC news announcements.

In this section, we discuss the stylised model of the rational expectations model as in Beaudry and Portier (2014) in order to explain how nonfundamentalness arise out of the presence of lagged shock effects of oil supply on the observables. To clarify the basic idea, consider the equilibrium endogenous variables, y_t, namely oil demand proxied by global economic activity, oil price and oil stocks, is equal to the sum of expected oil supply, p_t, as follows:

$$y_t = \sum_{j=0}^{\infty} \beta^j \mathbb{E}_t \left(p_{t+i} \right) + \nu_t, \;\; \beta < 1 \tag{3.1}$$

The dynamics for the oil supply can be assumed as follows:

$$p_t = \alpha p_{t-1} + u_{t-k} + \chi u_t \tag{3.2}$$

where u_t and ν_t are disturbances and $\mid \alpha \mid < 1$. In particular,u_t is a shock that has only an effect on oil production with a *k*-period lag, though being observed at time *t* by agents. This implies that agents will adapt their behavior already at time *t*, as they anticipate the future effects of u_t on oil supply already at time *t*. In other words,u_t enters the information set on which agents form their expectations for p_{t+k}. However, the econometrician will only be able to see the impact of u_t at time $t+k$. If $\chi < 1$, u_t at has a greater contribution to u_{t+q} than to u_t, in which case the process for p_t is primarily driven by shocks observed by the economic agents before t. Substituting the expression for $\mathbb{E}_t\left(p_t\right)$ into (2) leads to:

$$y_t = \frac{1}{1-\alpha\beta}p_t + \frac{\beta}{1-\alpha\beta}\left[\beta^{k-1}u_t + \cdots + \beta u_{t-k} + u_{t-k+1}\right] \tag{3.3}$$

The structural moving average representations of $(p_t, y_t)'$ is:

$$\begin{bmatrix} p_t \\ y_t \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ \frac{\alpha}{1-\alpha\beta} & 0 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ y_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \chi + L^k & 0 \\ \frac{\chi+\beta^k}{1-\alpha\beta} + \frac{1}{1-\alpha\beta}\sum_{i=1}^{k-1}\beta^{k-i}L^i + \frac{1}{1-\alpha\beta}L^k & 1 \end{bmatrix}}_{=A(L)} \begin{bmatrix} u_t \\ \nu_t \end{bmatrix} \tag{3.4}$$

where $(p_t, y_t)'$-fundamental process requires that $|A(L)| \neq 0$. However, as it can be easily showed from the simple case where k = 2, and $\chi < 1$ that $|A(L)| = 0$ for $L = \pm\sqrt{\chi}^i$.

In this case, the process has a nonfundamental representation MA due to roots in the unit circle. In other words, as soon as the news shock contributes to p_t relatively more, $\chi < 1$, the observable suffer from nonfundamentalness and no causal VAR representation in terms of $(p_t, y_t)'$ exists for structural shocks. In order to cope with the nonfundamentalness problem, one solution requires the use of non-causal VAR. As showed in Nelimarkka (2017a), the above nonfundamental representation can be mapped in Non-Causal VAR.

3.3 Non-Causal VAR

Within the purely autoregressive VAR model and in its MA representation, the vector of observable variables y_t is driven only by the present and past of the vector process of structural shocks which are fundamental by construction. Nonfundamentalness is, therefore, ruled out by the structural VAR model even though it may arise in a case when forward-looking agents have expectation about future values of the model variables. It is self-evident now from many theoretical macroeconomic models (e.g., Susan Yang, 2005; Leeper et al., 2013), that impulse response functions should be derived from non-fundamental representations. However, as it is pointed out by Lippi and Reichlin (1994), the identification of the structural shocks become more involved than with VAR model as many nonfundamental representations are possible. This gives rise to further problem related to choice of the right representation based on the economic theory.

Nonfundamental representations are related to Non-Causal VAR (NC-VAR) models (e.g., Lanne and Saikkonen, 2013a; Lanne and Luoto, 2016; Gourieroux and Jasiak, 2017a; Davis and Song, 2020; Nelimarkka, 2017b,a). As proved in Nelimarkka (2017a), within a general theoretical model, that under the hypothesis of nonfundamentalness due to the presence of a forward-looking behavior generates non-causality. These models explicitly account for the possibility of nonfundamental shocks which are represented by past and future terms.

Consider the following Bayesian of the $n-$dimensional NC-VAR(r,s) model developed by Lanne and Luoto (2016), where y_t depending on past and future observations is generated by:

$$\Pi\left(L\right)\Phi\left(L^{-1}\right)y_t=\epsilon_t \tag{3.5}$$

where the causal polynomial $\Pi\left(L\right)=I_n-\Pi_1 L-\cdots-\Pi_r L^r$, the non-causal polynomial $\Phi\left(L^{-1}\right)=I_n-\Phi_1 L^{-1}-\cdots-\Phi_s L^{-s}$ and ϵ_t is a sequence of independent, identically distributed (continuous) random vectors with zero mean and finite positive definite covariance matrix. L is a backward shift operator. Stationarity of the process and the existence of an MA representation is guaranteed by the assumption that the matrix polynomials $\Pi\left(L\right)$ and $\Phi\left(z\right)\forall z\in\mathbf{C}$ have their zeroes outside the unit disc, i.e. det$\Pi\left(z\right)\neq 0$ for $|z|\leq 1$ and det$\Phi\left(z\right)\neq 0$ for $|z|\leq 1$

We can write the previous process as one-sided moving average representation (MA):

$$\Phi\left(L^{-1}\right) y_t = \Pi\left(L\right)^{-1} \epsilon_t = M(L)\epsilon_t = \sum_{j=0}^{\infty} M_j \epsilon_{t-j} \tag{3.6}$$

It is more easily to highlight the forward-looking component by rewriting equation 3.6:

$$y_t = \phi_1 y_{t+1} + \cdots + \phi_s y_{t+s} + \sum_{j=0}^{\infty} M_j \epsilon_{t-j} \tag{3.7}$$

When lead terms ϕ_i are different from zero, indicating that y_t are non-fundamental as lags are insufficient to recover the structural shocks.

Finally, a two-sided MA representation of equation 3.6 can be written as follows:

$$y_t = \sum_{-\infty}^{\infty} \Psi_j \epsilon_{t-j}, \ \forall z, \ \Psi\left(z\right) = \Phi\left(z^{-1}\right)^{-1} \Pi\left(z\right)^{-1} \tag{3.8}$$

It is important to emphasize that the error term of the model ϵ_t is necessarily non-Gaussian, as otherwise, estimating noncausal NC-VAR(r,s) becomes equivalent to estimating a causal VAR($r + s$). More precisely, a multivariate t-distribution is assigned to ϵ_t as follows:

$$\epsilon_t = \omega_t^{-\frac{1}{2}} \eta_t, \tag{3.9}$$

where $\eta_t \sim$ N$(0, \Sigma)$ and $\lambda\omega_t$ is χ^2_λ-distributed, and $\omega_t^{-\frac{1}{2}}$ is the scalar volatility factor with λ degree-of-freedom. Consequently, variables in y_t can be characterised by fat tails for smaller λ or points towards normality when λ is substantially high.

3.3.1 Oil supply news shock identification

To implement our identification, the reduced-form NC-VAR specification, described in the previous section, consists of four endogenous variables, which are standard in the oil market empirical literature, in order: global oil production, global real economic activity, real oil price and global oil stocks. To recover structural shocks we need to impose some identification restrictions to the reduced-form error term implied by equation 3.9:

$$\epsilon_t = \omega_t^{-\frac{1}{2}} \eta_t = \bar{B} \bar{u}_t \tag{3.10}$$

where the t-distributed structural shock vector, $\bar{u}_t = \omega_t^{-\frac{1}{2}} u_t^* = [\bar{u}_{1,t} \cdots \bar{u}_{k,t}]' \sim t_\lambda (I_k)$ is a product of two latent factors, a k-dimensional vector of Gaussian shocks $u_t^* \sim N(0, I_k)$ and the volatility term $\omega_t^{-\frac{1}{2}}$.

Denote by W in equation 3.10 an orthogonal $(k \times k)$ matrix with w_i on its ith column, and $\tilde{A}$ a Cholesky factor such that $\Sigma = \tilde{A}\tilde{A}'$. The rotation matrix $\bar{B}$ is thus now given by $\bar{B} = \tilde{A}W$. As global oil production is ordered first in the model, so are oil supply news shocks $\bar{u}_{1,t}$ in $\bar{u}_t$. Impulse responses of global oil production are given by:

$$y_{1,t} = y_{1,t}^1 + \cdots + y_{1,t}^k, \;\; y_{1,t}^k = \sum_{j=-\infty}^{\infty} e_1' \Psi_j \tilde{A} w_l \bar{u}_{t-j} \tag{3.11}$$

with the vector $e_i = [0 \cdots 1 \cdots 0]'$ having one in its ith element and $W\bar{u}_{t-j} = w_1 \bar{u}_{1,t} + \cdots + w_k \bar{u}_{k,t}$. The identification procedure consists on solving for w_1 such that $\gamma_1 = \tilde{A} w_1$. Thus, we apply the max-share approach which require w_1 to maximise the share of unconditional variance of y_t driven by the news shock $\bar{u_{1,t}}$ over a time finite horizon $[H_1, H_2]$[1]:

$$\Omega_{1,y_{1,t}}^{[H_1,H_2]} = \frac{\mathbb{E}\left[\sum_{j=H_1}^{H_2} e_1' \Psi_j \tilde{A} w_1 \bar{u}_{1,t-j} \bar{u}_{1,t-j}' w_1' \tilde{A} \Psi_j' e_1\right]}{\mathbb{E}\left[\sum_{j=H_1}^{H_2} e_1' \Psi_j \Sigma \Psi_j' e_1\right]} = \frac{\sum_{j=H_1}^{H_2} e_1' \Psi_j \tilde{A} w_1 w_1' \tilde{A}' \Psi_j' e_1}{\sum_{j=-\infty}^{\infty} e_1' \Psi_j \Sigma \Psi_j' e_1} \tag{3.12}$$

This medium-run identification strategy is essentially the same as the *Max-Share* identification proposed by Uhlig (2004) and Francis et al. (2014) to measure technology news shocks, with a difference that, instead of maximizing the forecast error variance share, we follow Nelimarkka (2017a) and look for the shock that accounts for the maximum of the unconditional variance share explained by one shock for a finite time horizon H_1 and H_2. In this sens, oil supply news shocks are supposed to have an impact effect on forward-looking variables and a

[1]See Nelimarkka (2017a) for more details about the identification problem-solving where the maximisation problem can be resolved as an eigenvalue problem.

persistent but only medium-term effect on oil supply. To separate the oil supply news shock from the remaining disturbances, we therefore assume that the former makes the largest contribution to the unconditional variance of global oil production over an horizon window. We discuss in the next section our choice of the finite truncation horizon H_1 and H_2. Moreover, we allow the news shock to impact global oil production without delay.[2] As explained in Kurmann and Sims (2020), this leads to a more identification robustness to measurement errors and revisions in global oil production.

3.3.2 Data and Estimation

We estimate a monthly NC-VAR models using the standard global oil market endogenous variables mentioned above:

$$y_t = [\Delta \textit{prod}_t, \textit{grea}_t, \textit{rpo}_t, \Delta inv_t]$$

where $\Delta \textit{prod}_t$ denotes the percentage change in global oil production, obtained from the US Energy Information Administration's Monthly Energy Review, $\textit{grea}_t$ the global real economic activity proxied either by the last version of Kilian (2009)'s index reported in Kilian (2019) or the growth rate of OECD+6 industrial production proposed by Baumeister and Hamilton (2019a).[3] $\textit{rpo}_t$ the real oil price measured by the refiner acquisition cost (RAC) for imported crude oil and deflated by the US consumer price index, and Δinv_t the proxy for the change on global oil stocks constructed as in Kilian and Murphy (2014b). Estimation data span the period between 1973:02 and 2019:12.

The NC-VAR(r,s) models are estimated by employing a lag and lead length of 12 in order to capture the full dynamic of observables. As well explained in Nelimarkka (2017a), this choice

[2]Note also that, as explained in Nelimarkka (2017a), impact effect restriction is difficult to implement within the two-sided MA representation of the NC-VAR model.

[3]As proxy for the global real economic activity, we also used the new global economic activity index (GECON) constructed by Baumeister et al. (2020) and available from the Baumeister's website. We also use the updated and corrected version of the index of global real economic activity in industrial commodity markets from (Kilian, 2009, 2019) which is available from the Kilian's website. Results are reported in Figure 3.11 for GECON and in Figure 3.12 for Kilian's index, both in Appendix 3.9.

permits for observables to fully catch structural innovations in case when the NC-VAR(r,s) reduces to a standard autoregressive VAR in absence of non-fundamentalness problem, and otherwise for the non-causal part to have a rich structure if non-fundamental problem arises particularly because of anticipation.

Moreover, the NC-VAR is estimated with Bayesian methods as described in Lanne and Luoto (2016). To define priors for the NC-VAR dynamic parameters, we follow Nelimarkka (2017a) and impose a standard Minnesota prior. Specifically, hyperparameters that control the tightness of the priors on lags and leads are chosen so that variable dynamics are more driven by lag coefficients as lead coefficients are shrunken more heavily toward zero.[4] This helps attaining the unimodalilty when posterior distributions are more likely to be multimodal when estimating the NC-VAR model (Lanne and Luoto, 2016). Then, a numerical Bayesian approach (Gibbs sampling) is used to estimating NC-VAR models. Results are based on 10,000 posterior draws obtained after a burn-in of length 50,000. The recursive means of the retained draws (reported in appendix 3.7) show little variation thus providing evidence in favour of the convergence of the samplers.

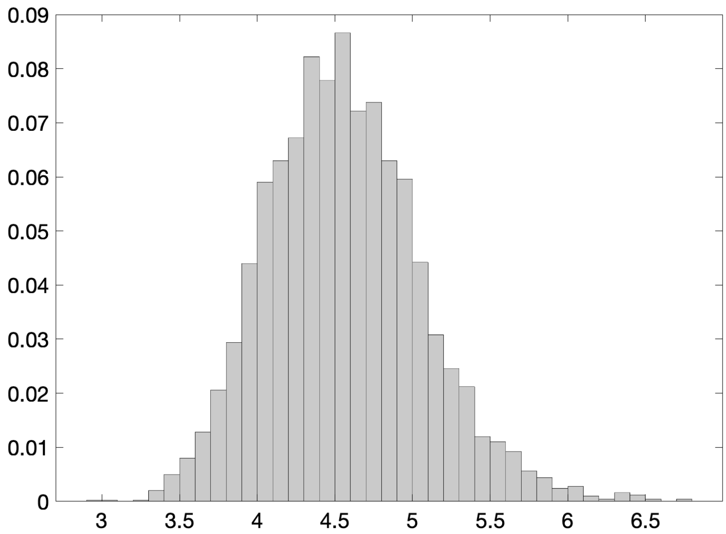
3.4 Empirical Results

In this section we present the results of impulse response function analysis to the oil supply news shock. Recall that the key advantage of the NC-VAR model is that it presents an attractive solution to the problem of nonfundamentalness. Similarly and following Nelimarkka (2017a) our identification of the news shock is carried out within the the NC-VAR model without the need to external additional *proxy*. Oil supply news shock is therefore the shock that contribute the most to global oil production variations at long but finite horizon.

Our first step is to check whether a nonfundamental representation is supported by the data. As pointed out by Lanne and Saikkonen (2013a), Gourieroux and Jasiak (2017a) and

[4]We set overall tightness prior value equals to 0.5 and decay parameter prior value equals to 1. For the tightness parameters regarding the lag coefficients, we choose 0.2 and for the lead coefficient 0.15. For more details about prior information, reader can refer to Appendix pages 28-29 in Nelimarkka (2017a).

Figure 3.1: Posterior density of DOF parameter λ

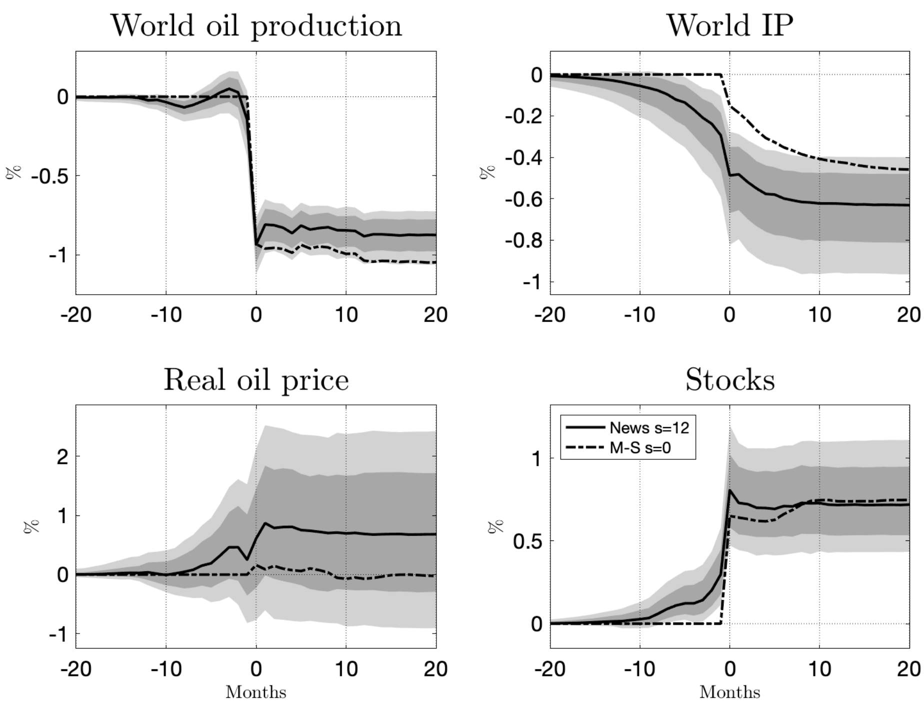


Notes: gray bars represent the frequency distribution of the DOF parameters from the 3-variable NC-VAR(12,12). The red line is the posterior mean

Davis and Song (2020) the validity of the non-Gaussian assumption of the error term is required justifying nonfundamental representation. Figure 3.1 plots the estimated marginal posterior density of the degrees of freedom (DOF) λ. The histogram indicates strong evidence in favour of fat tails as the posterior density is centred around 4.5 degree of freedom with the probability of λ being greater than 6 is almost nil.[5] Similarly, distributions of the residuals estimated by causal BVAR and densities of the observed variables, showed in Figure 3.4 in Appendix 3.6, point towards non-normality. This suggests that the normality assumption is inappropriate, confirming the choice of multivariate t-distribution for the error term, making it possible identifying a unique NC-VAR(r,s) specification. This indicates therefore that capturing the non-causal component is crucial for the data we study.

[5] As in Nelimarkka (2017a), we set the prior mean of λ to 10. With a very low posterior mean of λ (less than 5), data dominate the assumed prior mean of λ.

Figure 3.2: Impulse-response functions to news shock for Baumeister and Hamilton (2019a) 4-variable model



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister and Hamilton (2019a). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

3.4.1 How does the oil supply news shock diffuse to the oil market variables?

The shock of interest in the model is the first shock, namely the oil supply news shock. Figure 3.2 shows the cumulative impulse responses to one standard-deviation oil supply news shock (solid lines) along with, for a comparison purpose, responses to oil supply shock from VAR model estimation with MS identification (dashed lines). Oil news supply shock is normalized to represent a news shock of oil supply shortfall. Light and dark shaded bands represent 90% and 68% posterior credible sets, respectively. In the non-causal model context, the left side of the x-axis is added to represent responses related to the leads terms of the MA representation. The news shock is measured with a truncation horizon of $[H_1 = -12, H_2 = 12]$, but a similar results, reported in Figures in Appendix 3.8, are obtained with shorter and longer truncation horizons.

There is a clear evidence from the estimated impulse responses that a non negligible part of the decline on oil supply triggered by the news shock can be anticipated before the drop materializes. The median response estimates from the non-causal model suggest a large and significant responses of non-causal components for most variables. As our news shock identification allows global oil supply to react to the shock at its leads and on impact, this shock has a temporary and barely significant decrease on global oil production at the anticipatory lags. At these negative lags, global economic activity reacts negatively and significantly about four months ahead, revealing the forward-looking nature of this variable. Simultaneously, stocks start to increase significantly ahead of the drop of global oil production. Surprisingly, real oil price increases only modestly with a statistically insignificant reaction. As for the causal part, global oil supply, global economic activity and stocks react significantly on impact. Then, the estimated shock identifies persistent and significant decline on both global oil production and global economic activity, when it permanently increases stocks. Reactions of global oil production and global economic activity are different from those found by Känzig (2019) in that these variables react on impact to the news shock. This can be explained on one hand by our shock identification strategy, and on the other hand by the inclusion of lead terms in NC-VAR, allowing shock to be anticipated, resulting on correctly impact effects under nonfundamentalness. Regarding the oil stocks reaction, it appears consistent with the view that when market players expects a shortfall in the future, they increase their demand of oil as a precautionary measure building up inventories today. The oil price permanently increases but statistically insignificantly.

Overall, our identified shock induces persistent changes with a smooth anticipation effect in global oil production. This is consistent with the view that oil supply news leads to a fall in global oil production, with a distinction that this later reacts without a delay to the shock. This shock also significantly impacts forward-looking variables, namely global economic activity and stocks, but not on real oil price before it materialises in global oil production. This evidence suggests that the oil supply news shock can be viewed as a strong signal from the future oil supply and a trigger of short-run fluctuations on global economic activity and stocks but in much lesser manner to real oil price. The likely reason for this finding is that, as pointed out by Kilian and Zhou (2020), the extent that supply news matter for real oil price depends on expectations of oil demand. It is clear from impulse response estimates that the reaction of global economic activity results in partially offsetting the reaction of real oil price, despite the increase in precautionary demand.

For a comparison purpose, the estimated impulse response to the news shock based on the Max-Share identification scheme is shown in dashed lines in Figure 3.2. Although median estimated responses of all variables, at the positive side of the horizon, have fairly the same patterns, they differ in terms of magnitude for some variables. While median responses of global oil production and stocks are fairly similar, those of global economic activity and real oil price are larger when using NC-VAR.

Sensitivity checks: As already mentioned, using alternative measures of the global economic activity, namely Kilian (2009)'s index and that recently proposed by Baumeister et al. (2020), leads to similar results shown in figures 3.11 and 3.12 in Appendix 3.9. In addition, we show that the results survive if a different truncation windows, with different value of H_1 and H_2, are used to identify news shock in the baseline model(see figures 3.8 in Appendix 3.8). In addition, as we show in the next section, our results are also robust to the inclusion of additional macroeconomic and financial variables to the system. Finally, it is evident that the transmission of the oil supply shocks can evolves depending on the production capacity in oil producing countries. This due to the temporal instability of capacity constraints. Particularly, the rapid and persistent increase in the real oil price during the period 2003-2008 can be explained by the inability of the production capacity due to an underinvestment to respond to increasing global oil demand, especially from emerging countries. This period coincides also with the increasing intervention of commodity index investors on futures oil market. This financialisation phenomenon could also have changed the oil price behaviour (Tang and Xiong, 2012; Charlot et al., 2016; Juvenal and Petrella, 2015, among other). Given this evidence, we estimate our model using sample that span 1973:1 to 2004:12. As shown in Figure 3.13 in Appendix 3.10, truncating the sample to December 2004 does not affect any of the variable reactions, providing qualitative support for the benchmark estimates.

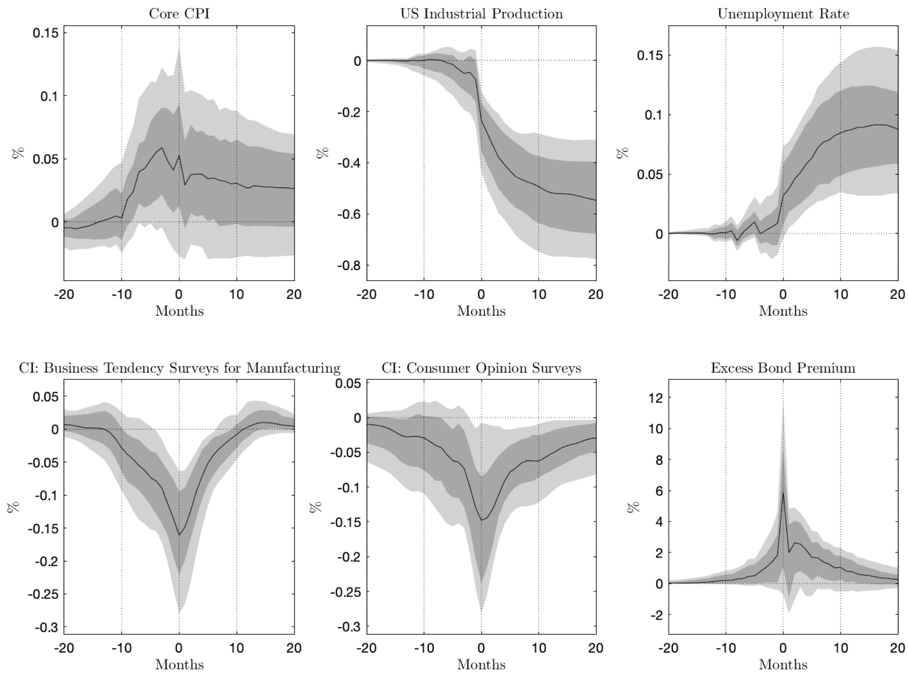
3.4.2 What impact does the news shock have on macroeconomic variables?

To address this question, additional variables are included in the system, one variable at a time in order to avoid to estimate a large model: US core CPI, US industrial production, US unemployment rate, US business confidence indicator, US consumer confidence indicator and the excess bond premium (EBP).[6]

Figure 3.3 shows that while core CPI, consumer confidence indicator, business confidence and EBP anticipate the news shock about oil supply shortfalls, industrial production and unemployment rate do not react during the anticipatory lags. The news shock impacts positively core CPI and EBP, but barely significantly. However, consumer and business confidence indicators decline significantly indicating that oil supply news shocks can be expected by economic agents. At the causal part, core CPI continues to rise but insignificantly. The US industrial production reacts contemporaneously and significantly to the shock with a similar pattern as that of global economic activity and simultaneously US unemployment rate increases significantly. Noteworthy that consumer and business confidence indicators decrease significantly on impact and start to recover and return to their initial values after about eight months. Interestingly, the EBP, although statistically insignificant, peaks on impact and start to converge to its initial value, pointing to a deterioration of the economic activity. These findings suggest, therefore, that oil supply news shock can have large effects on the macroeconomy, and in particular news about future oil supply shortfalls can have recessionary effects.

[6]All variables are originated from FRED database, except for excess bond premium which is an updated version of the measure of Gilchrist and Zakrajšek (2012) available from the Fed website: https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/updating-the-recession-risk-and-the-excess-bond-premium-20161006.html

Figure 3.3: Impulse-response functions to news shock



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister and Hamilton (2019a). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

3.5 Conclusion

Purely autoregressive VAR models have become prominent in empirical works on global oil market in spite of the presence of nonfundamentalness problem which may arise as result of the informational insufficiency in the small-scale VAR model. Dealing with this issue is crucial so that reduced form innovations contain enough information to recover the structural shocks of interest.

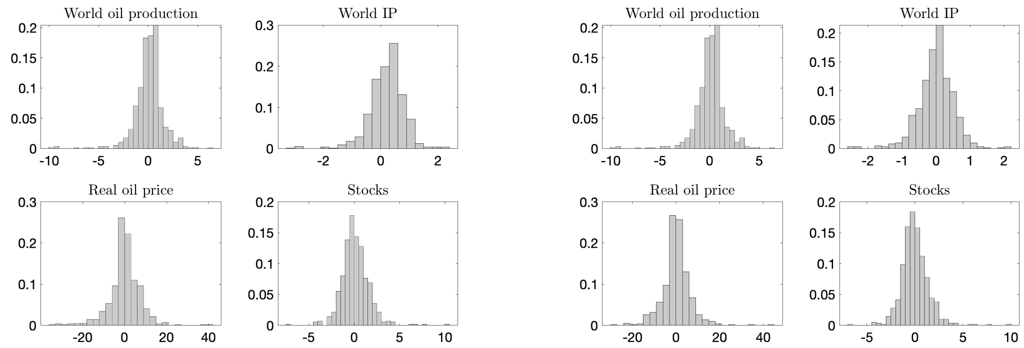
until now, nonfundamentalness issue on the global oil market has been addressed by either augmenting small-scale VAR models by additional variables or latent factors, or using external instrument or proxies leading to more credible identification scheme. In this paper we dealt with this issue by employing the non-causal VAR model on standard global oil market, namely global oil production, global economic activity, oil stocks and real oil price and US macroeconomic variables in order to analyse the effect of the oil supply news shock.

We showed, first, that the nonfundamental representation is supported by the data, justifying therefore the use of Non-Causal VAR as an option dealing with the information deficiencies when modelling global oil market. Second, follow Nelimarkka (2017a) we identified an oil supply news shock as a shock that drives global oil production the most for a finite time horizon. We further showed that our identified oil supply news shock is anticipated by forward-looking variables before the shock materialises, highlighting the prominent role of expectations in propagating the shock. We documented also that a negative oil supply news shock results in abrupt and permanent reaction in global oil production, global economic activity and in oil inventory. However, the oil supply shock has only a limited effect on oil price. Moreover, news shock about oil supply shortfalls do have macroeconomic consequences as it causes a substantial decline in US industrial production. Finally, evidence on the prominent role of the expectation channel is confirmed by the reaction of consumer and business confidence indicators over the anticipatory lags before the shock materialises.

While NC-VAR represents a promising approach dealing with nonfundamentalness problem, the issue of identification of structural shocks remains challenging task. The lack of development of structural NC-VAR is unfortunate, preventing the analysis of the effect of further interesting shocks using credible identification scheme. As further research, it would be worthwhile to explore the impact of the oil demand shock on global oil market and macroeconomic variables.

3.6 Non-normality

Figure 3.4: Histograms of Baumeister and Hamilton (2019a) 4-variable and the associated residuals from causal VAR(12) estimation

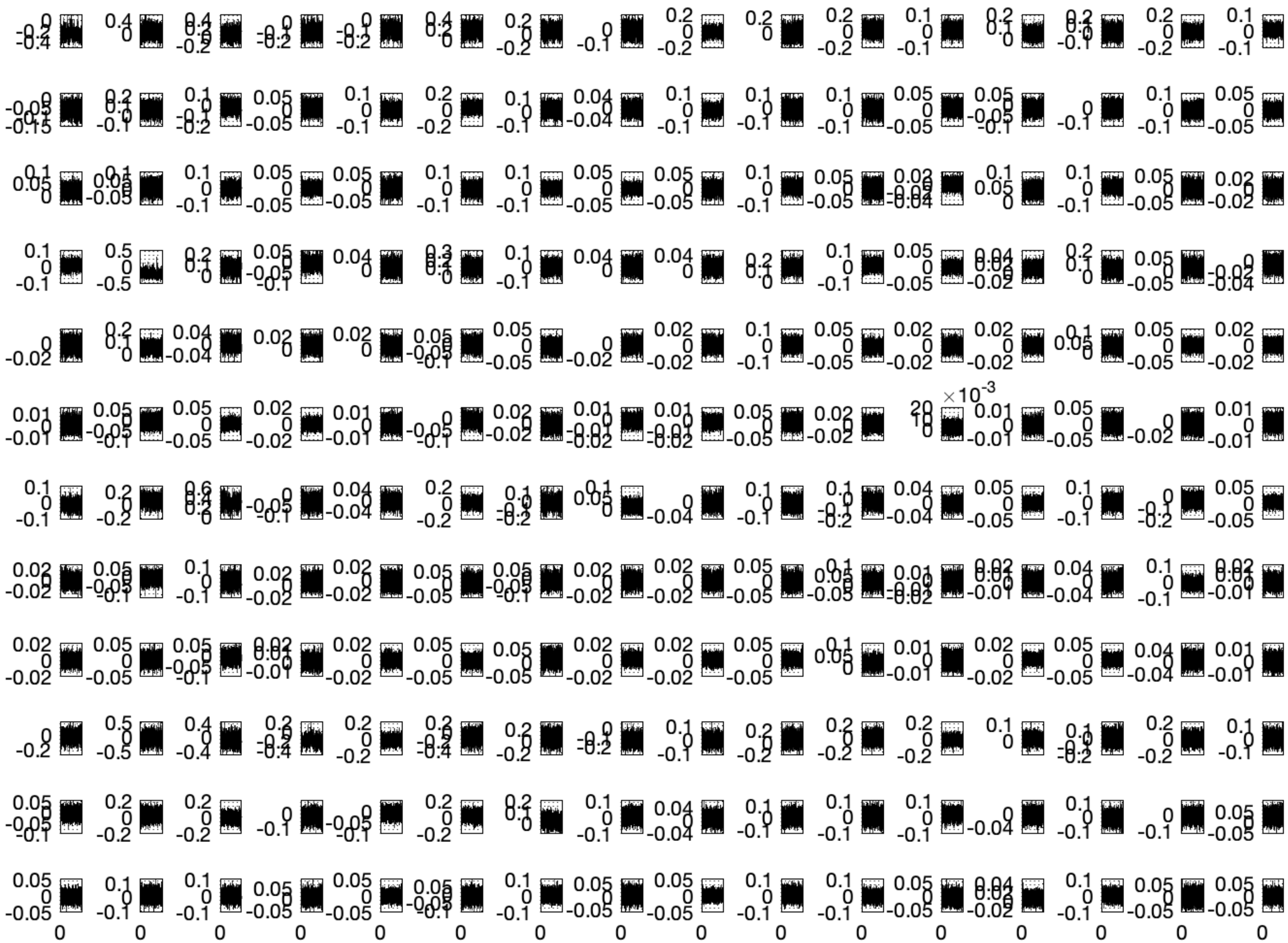


(a) Variables

(b) VAR(12) residuals

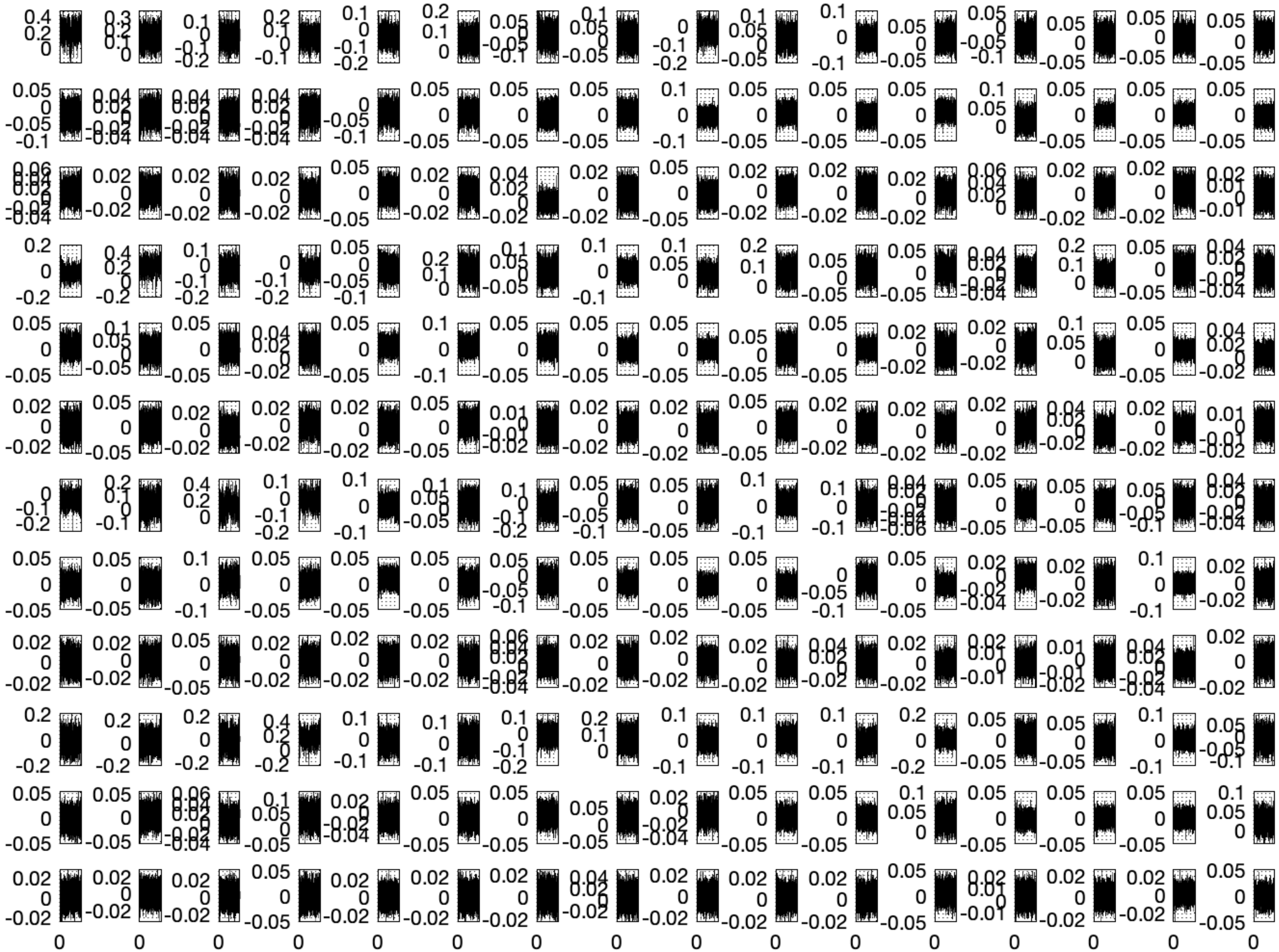
3.7 Convergence of the posterior sampler

Figure 3.5: Paths of the Markov chains for the draws of elements in ϕ of Baumeister and Hamilton (2019a) 4-variable model noncausal VAR(12,12)



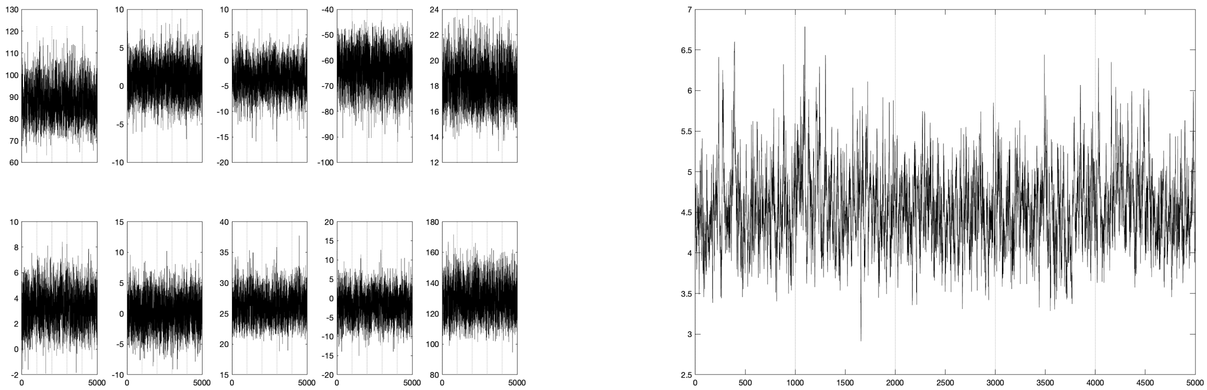
Notes: The x-axes correspond to the draws, the y-axes to the parameter values

Figure 3.6: Paths of the Markov chains for the draws of elements in π of the Baumeister and Hamilton (2019a) 4-variable model noncausal VAR(12,12)



Notes: The x-axes correspond to the draws, the y-axes to the parameter values

Figure 3.7: Paths of the Markov chains for the draws of scale matrix and the degrees-of-freedom parameter of Baumeister and Hamilton (2019a) 4-variable model noncausal VAR(4,4) model



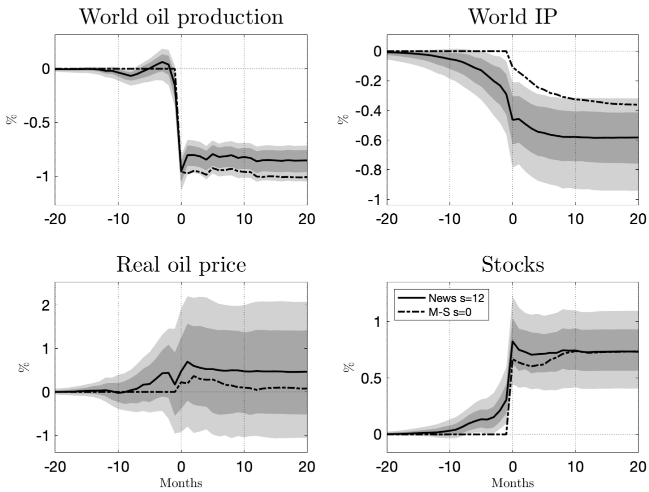
(a) vech(Σ)

(b) λ

Notes: The x-axes correspond to the draws, the y-axes to the parameter values

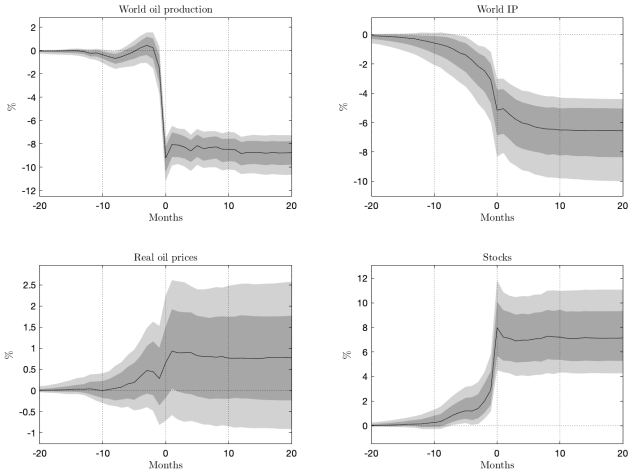
3.8 News shock identification with different truncation windows (H_1; H_2) robustness

Figure 3.8: Impulse-response functions to news shock for Baumeister and Hamilton (2019a) 4-variable model: H_1=-6 and H_2=6



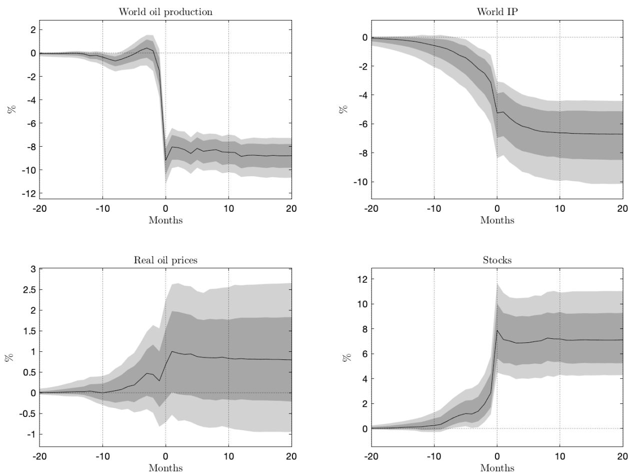
Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister and Hamilton (2019a). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

Figure 3.9: Impulse-response functions to news shock for Baumeister and Hamilton (2019a) 4-variable model: H_1=-12 and H_2=24



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister and Hamilton (2019a). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

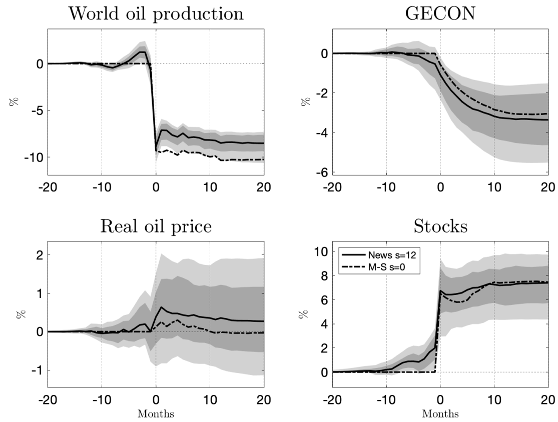
Figure 3.10: Impulse-response functions to news shock for Baumeister and Hamilton (2019a) 4-variable model: H_1=-12 and H_2=60



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister and Hamilton (2019a). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

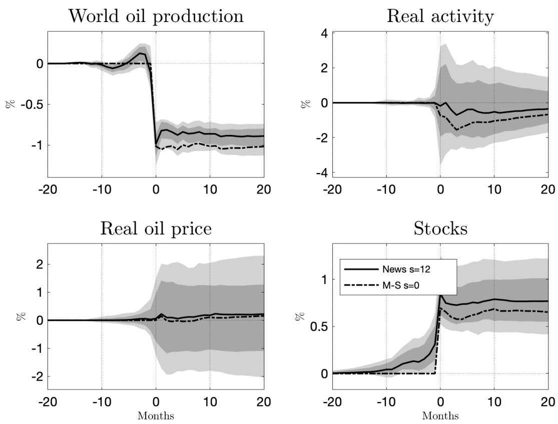
3.9 Baseline NC-VAR model using different global economic activity measures

Figure 3.11: Impulse-response functions using GECON index



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister et al. (2020). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

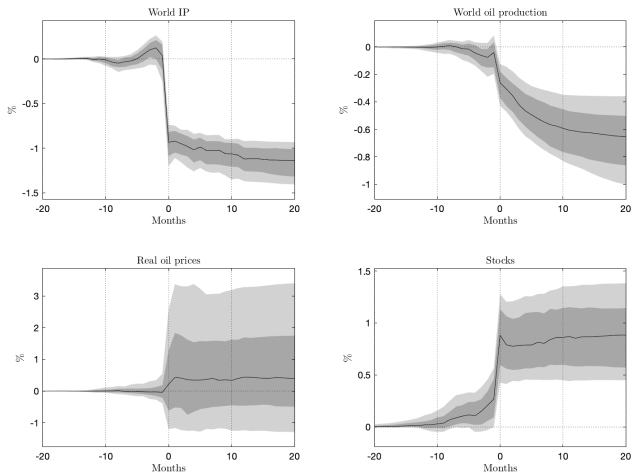
Figure 3.12: Impulse-response functions to news shock using Kilian (2009)'s index



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister et al. (2020). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

3.10 Before 2005

Figure 3.13: Impulse-response functions to news shock for Baumeister and Hamilton (2019a) 4-variable model: H_1=-12 and H_2=12 / 2005



Notes: The black dashed lines are the posterior median responses of the 4-variable baseline model from Baumeister and Hamilton (2019a). The solid lines are the posterior median responses of the NC-VAR(12,12). Both models are identified using Nelimarkka (2017a) news technique. All responses are shown in levels. Light and dark grey shaded regions are the 90 % and 68 % credible sets of the NC-VAR(12,12). Because of the noncausality, the impulse responses are located in both sides of zero. The negative side corresponds to the lead terms of the MA representation of NC-VAR.

Chapter 4

The role of expectations in predicting the real prices of oil: a non-causal analysis

[This chapter is based on Thomas (2020)]

4.1 Introduction

The question of what determines the dynamics of oil spot and futures prices, and to a lesser extent our ability to predict it, is of constant interest. To answer this question, different strategy have been applied. The first one is based on the construction of structural models containing macroeconomic variables representative of the oil market (see e.g., Baumeister and Kilian, 2012, 2014; Baumeister et al., 2014; Baumeister and Kilian, 2015; Baumeister et al., 2017b, 2020). These models are more accurate than the no-change forecast.

The second strategy is derived from market efficiency theory. Alquist and Kilian (2010a) using a two-country, multi-period general equilibrium model of the spot and future of crude oil prices, in conjunction with empirical analysis, highlights the link between the spot price of crude oil, expectation of future oil prices, the prices of crude oil futures and the oil futures spread (defined as the deviation of futures prices from the spot future prices). Nevertheless, Alquist and Kilian (2010a) empirical main results is that the oil futures prices fail to improve the accuracy against simple non-change forecasts.[1] This contradicts widely held views among policymakers and financial analysts. The result was a robust finding that the no change forecasts is much more accurate across all horizons from 1 month to 12 months for spot prices in level. Alquist and Kilian (2010a) shows that the variability of the future prices about the spot prices is captured by the spread of oil futures. This leads to the second result, which is the existence of a convenience yield for crude oil. This existence is derived from the use inventories to avoid the outages of the production process or to match unexpected shifts in demand. This "option value" is reflected in convenience yield (see e.g. Brennan and Hughes, 1991; Pindyck, 1994; Schwartz, 1997; Routledge et al., 2000).

Based on the convenience yield, Schwartz (1997) assess the stochastic behavior of commodity prices with a variation of the two-factor Gibson and Schwartz (1990)'s model. The first factor is the logarithm of the spot price of the commodity which is assumed to follow a mean reverting process of the Ornstein-Uhlenbeck type. The second factor in this model is the convenience yield of the commodity and it is assumed to follow a mean reverting process. Schwartz (1997) assumes that the agents don't have expectations about future convenience yield values. Gospodinov and Ng (2013), shows that the convenience yields explain commodity prices and can be seen as informational variables about future economic conditions as conveyed by the futures markets. In this paper we try to release these assumption, based on the following results: First, rational expectations formed by agents, in asset pricing, lead to non-causal representation, particularly when forward-looking variables are involved(Lof, 2013; Lanne and Saikkonen, 2008, see e.g.,). Non-causal autoregression allows y_t to depend both on past and future observations. Second there is a recent growing literature in favor of noncausal processes, which seem to better fit commodity price series (see e.g., Lanne and Saikkonen, 2013b; Gourieroux and Jasiak, 2017b). Last but not least, non-causal autoregression in univariate case is solution of the stochastic equation that drives mean reverting process of the

[1] Econometric approaches are based on the futures prices of oil, and a professional survey approach

Ornstein-Uhlenbeck type (see e.g., Behme, 2011; Fries, 2018; Gourieroux et al., 2019). Extend these solution to Schwartz (1997)'s multivariate equations rise many theoretically hardships along the way.

Our idea is to investigate the predictive power of the convenience yield, by modifying the empirical specification to include agents expectations. We estimate bivariate Bayesian Noncausal VAR from Lanne and Luoto (2016).[2] Our endogenous variables are the real energy spot prices and the associated convenience yield. We also perform a pseudo-recursive exercise to assess the forecasting performance of our specification.

Our contribution is twofold. First, we provide evidence that expectations parameters between spot prices and convenience yield are highly significant thereby indicating the empirical relevancy of our specification to capture markets expectations formed by markets player in oil markets. Second, as our approach significantly reduces the number of estimated parameters compared to the reference models in the literature, it is likely to perform well in a forecasting exercise. Accordingly, we provide empirical evidence that real-time forecasts of real oil prices can be notably more accurate than the no-change forecast but also much more accurate than real-time forecasts generated by existing structural models relying on either frequentist or bayesian VARs as in, among others (Baumeister and Kilian, 2012, 2014). Real-time mean squared prediction error (MSPE) reductions may be as high as 23% 1 month ahead and 29% 12 months ahead.

Our methodology could be applied to different commodity as long as they are storable and the futures prices are available. For robustness, and to challenge our approach we decide to extend our empirical exercise to U.S. gas market (Knetsch, 2007). The comparison between the two is interesting, indeed gas is a local market with a large storage and transport capacity but expensive in infrastructure costs and much less flexible than oil which is a global market with low storage and transport costs (Rosendahl and Sagen, 2009). We find that our specification is also relevant for U.S. natural gas and our real-time forecast is notably more accurate than

[2] Our idea is that the VAR(1,1) process may be a discrete time approximation of the anticipated part of the Schwartz (1997)'s 2-factor model.

the no-change.

Beyond the traditional analysis at the monthly frequency for which expectations are likely to play a role based on inventories adjustments, we further investigate the forecasting accuracy of our model at the daily and weekly frequency. We are going further at these frequencies for two reasons: we first investigate if agents expectations are similarly formed in higher frequencies, second we became closer to the use case Schwartz (1997)'s models for option prices purposes. Forecasts are still better than the no-change but the gain is inferior than at the monthly frequency. This, nevertheless, is good news for some potential implementation of investment strategies that could rely on predictions based on the convenience yield.

The remainder of the paper is as follows. The next section introduces the underlying theory from Schwartz (1997) and research advancement on theoretical level. Section 4.3 introduces the econometric framework. We briefly review the Bayesian non-causal vector autoregression methodology and the associated estimation and forecasting procedure from Lanne and Luoto (2016). In Section 4.4, we present empirical estimation of expectations between the spot prices and convenience yield and we provide new empirical evidence of predictive power of convenience yield for the oil market. Next we extend our approach to natural gas, and for different data frequency, to prove the relevancy and the robustness of our empirical specification. Finally, Section provides concluding remarks

4.2 Underlying theory

This section is an ongoing work, we present the current state of our research on this subject, and propose a research plan. Our idea is to show that the VAR(0,1) process could be a discrete time approximation of the anticipated part of the Schwartz (1997)'s 2-factor model. These models are some of the most widely used for the evaluation of option values. The first factor is the spot price S_t and the second is the convenience yield δ_t. These two factors are defined by the following joint process:

$$dS_t = (\mu - \delta)Sdt + \sigma_1 Sdz_1 \tag{4.1}$$

$$d\delta_t = \kappa(\mu - \delta)dt + \sigma_2 dz_2 \tag{4.2}$$

where dz_1 and dz_2 are the increments of two Brownian motion which are correlated as follows: $\mathbb{E}(dz_1, dz_2) = \rho$. Equation 4.2 is the definition of an Ornstein-Uhlenbeck (OU) process.

Our current theoretical investigation leads to: AR(0,1) is a solution of equation 4.2 with Levy noise (Behme, 2011) or with noise generated by α-stable laws (Fries, 2018; Gouriéroux and Lu, 2019). Our first idea is to extend the solutions of (Behme, 2011; Gouriéroux and Lu, 2019) to shows that AR(0,1) is also a solution with a t-distributed noise. One key issue is the fact that Student's laws are not include in the class of stable probability laws. To do that we try to prove the existence of a linear approximation of the non-causal part AR(0,1) in the case of a t-distribution using the modeling of Barndorff-Nielsen and Shephard (2001); Heyde and Leonenko (2005), numerically and empirically. These univariate, Non-Gaussian OU process have already been the subject of a publication on commodities (Chevallier and Goutte, 2017), which proves the interest of this modeling in our framework. Then the univariate process must be extended to the VAR process. This is a not a major concern as in (Gouriéroux and Lu, 2019). To replicate Schwartz (1997) models our issue comes from the correlation of the two processes $dz_1 dz_2 = \rho$. One of the solution that we are currently experimenting is the use of a density modeling joined with a t-Student's Copula (Kallsen and Tankov, 2006). This has a double advantage, it allows to tackle Gaussian-Dependence and to ensure stability in law. Its also facilitate the generation of sample from the so-called t-distributed Schwartz (1997)'s models. For option pricing another key problem that we identified, is the use in our previous model of a non zero value of both purely causal and purely non-causal parts. Gouriéroux and Lu (2019), clearly demonstrate that, if the process have a non zeros representation of causal parts, the proprieties in reverse time and in calendar time are not equal. Two applications naturally follow from this part: First the implementation of new option evaluation methods that could be extended to many different commodities, and second, definition of new closed forms for options and future prices of the 2-factors (Schwartz, 1997)'s model modified to include expectation formation mechanism. This theoretical works is empirically justified by the results in the following sections.

4.3 Econometric framework

This section introduces the empirical specification that allows us to include expectations between spot prices and convenience yield. We briefly review the Bayesian non-causal vector autoregression methodology and the associated, estimation and forecasting procedure from Lanne and Luoto (2016).

4.3.1 Non-causal multivariate framework

The n-dimensional non-causal VAR(r,s) process is developed in Lanne and Luoto (2016). y_t is allow to depend on past and future observations:

$$\Pi\left(B\right)\Phi\left(B^{-1}\right)y_t = \epsilon_t \tag{4.3}$$

where $\Pi\left(B\right) = I_n - \Pi_1 B - \cdots - \Pi_r B^r$, $\Phi\left(B^{-1}\right) = I_n - \Phi_1 B^{-1} - \cdots - \Phi_r B^{-r}$ and B is a backward shift operator. ϵ_t is a sequence of independent, identically distributed (continuous) random vectors with zero mean and finite positive definite covariance matrix. It is important to emphasize that the error term of the model ϵ_t is necessarily non-Gaussian, as otherwise, estimating noncausal VAR(r,s) becomes equivalent to estimating a causal VAR($r+s$). The stationarity of the process is guaranteed by the assumption that the matrix polynomials $\Pi\left(B\right)$ and $\Phi\left(z\right)$ have their zeros outside the unit disc, i.e. det$[\Pi\left(z\right)] \neq 0$ for $|z| \leq 1$ and det$[\Phi\left(z\right)] \neq 0$. We can write the previous process through a moving average representation for both

$u_t = \Phi(B^{-1})y_t$ and $u_t = |\Pi(B)|\, y_t$ as follow:

$$u_t = \sum_{j=0}^{\infty} M_j \epsilon_{t-j} \tag{4.4}$$

$$w_t = \sum_{j=-(n-1)r}^{\infty} N_j \epsilon_{t+j} \tag{4.5}$$

Finally, a two-sided MA representation of equation 4.3 can be written as follows:

$$y_t = \sum_{-\infty}^{\infty} \Psi_j \epsilon_{t-j} \tag{4.6}$$

where $\forall z,\ \Psi\left(z\right) = \Phi\left(z^{-1}\right)^{-1} \Pi\left(z\right)^{-1}$. The following conditional expectation of the VAR(r,s) moving average representation highlights the process dependence on future errors.

$$y_t = \sum_{-\infty}^{s-1} \Psi_j \mathbb{E}\left(\epsilon_{t-j}\right) + \sum_{s}^{\infty} \Psi_j \epsilon_{t+j} \tag{4.7}$$

4.3.2 Estimation

Estimation method builds upon the work of Lanne and Saikkonen (2013b) and on the previous paper on the Bayesian analysis of noncausal AR models from Lanne et al. (2012). In particular, our basic estimation algorithm is a straightforward extension of their Metropolis-within-Gibbs sampler. Algorithm estimation is based on an algorithm based on a Mixture of t-distribution by Importance Sampling weighted Expectation Maximization (MitISEM, Hoogerheide et al., 2012). In so doing we follow Cappé et al. (2008), who argue that non-causal autoregressions lead to a multi-modal density. In our case we use a mixture of multivariate t-distributions as the candidate density.

4.3.3 Forecasting

Let us now consider the evaluation of the posterior predictive distribution: y_{T+h}. We shall assume that the model is both non-causal and multivariate. We can approximate the infinite sum in Eq. 4.7 as:

$$|\Pi(B)| = a\left(z\right) = 1 - a_1 z - \cdots - a_{nr} z^{nr} \tag{4.8}$$

and by approximation, the infinite sum can be:

$$y_{t+h} \approx a_1 y_{T+h-1} + \cdots + a_{nr} y_{t+h-nr} + \sum_{j=-(n-1)r)}^{M-h} N_j \epsilon_{T+h+j} \tag{4.9}$$

From the previous equation, we observe that the forecasting procedure is based on the estimation of the joint density of the augmented data: $(\mathbf{y}, \epsilon_{t+1}, \cdots, \epsilon_{t+M})$. The simulation procedure for the augmented data density is an extension of the univariate one in Lanne et al. (2012). We estimate the posterior distribution of the parameters ϕ, π, Σ and the latent variable $\widetilde{\omega}$, ϵ^+, by using the joint (conditional) density of $\mathbf{y}$ and ϵ^+. Due to high-dimensional posterior distribution ϵ^+, we use the method of Chib and Ramamurthy (2010).

4.4 Empirical analysis

All the estimations and computations presented in this paper were carried out using Julia. As recursive estimation of different models can be very time consuming, we take advantage of the computing clusters through the use of parallel computing in Julia software. In addition, Julia is open source and free, thereby facilitating reproduction of the results.

4.4.1 Empirical framework

To test the endogenous anticipative behavior between the spot prices and convenience yield we specified y_t as follows:

$$y_t = [CY_t, RAC_t]$$

Refiner Acquisition Cost of Crude oil Price (RAC), according to the EIA website is defined as the cost of crude oil, including transportation and other fees paid by the refiner.[3] RAC is considered as a proxy for the price of Brent crude oil, which has become increasingly accepted as the best measure of the global price of oil in recent year (Baumeister and Kilian, 2014).

[3]The refiner acquisition cost does not include the cost of crude oil purchased for the Strategic Petroleum Reserve (SPR)

Forecasts of the real price of oil are routinely used by international organizations and central banks worldwide in assessing the global world price and domestic economic outlook (Baumeister and Kilian, 2014). RAC is only available at monthly frequency.

We construct convenience yield CY based on Working (1949)'s using LIBOR as the risk-free rate. We choose futures prices for West Texas Intermediate Crude oil to construct the convenience yield.[4] The data source is Bloomberg. We obtain a value of convenience yield, for maturity 1, 2, 3, 6 and 12 months-ahead based on the associated futures contracts. We choose to perform a Principal Component Analysis (PCA) over all maturities and include the first component of the PCA in our bivariate specification. In the case of oil, the first component explains 96% of the total variance. For CY, data are readily available and we see that the extension to higher frequencies (Weekly, Daily) is relatively easy. Estimation data span the period between 1989:06 and 2019:05.[5]

We set the prior hyperparameter $\underline{\lambda} = 5$ because small values of $\underline{\lambda}$ seem to increase the performance of the sampler (Lanne and Luoto, 2016). The VAR coefficients π and ϕ are assumed to be prior-independent. The elements of the hyperparameters $\underline{\phi}$ and $\underline{\pi}$ are set to zero. Following Litterman (1986), we set diagonal element $\underline{V_\pi}$ and $\underline{V_\phi}$ such that the prior standard deviations of the parameters for own and foreign lags (or leads) equal $\frac{\gamma_1}{l^{\gamma_3}}$ and $\sigma_i \gamma_1 \gamma 2/\sigma_j l^{\gamma_3}$ respectively, where $l = 1, \cdots, r$ (or $l = 1, \cdots, s$). $\gamma_1 = 2$, $\gamma_2 = 1$ and $\gamma_3 = 1$. σ_i^2 is set at the residual standard error of a univariate causal AR(p)(p=r+s) model for $i = 1, \cdots, n$. The degrees of freedom parameter $\underline{\nu} = 10$. We assume Σ is diagonal. $\underline{S} = \left(\underline{\mu} - n - 1\right)$ diag$(\sigma_1, \cdots, \sigma_n^2)$.

4.4.2 The role of anticipations

Non-causality raises problems in identifying the coefficients r and s. Indeed, Lanne and Luoto (2016) shows the fact that it's impossible, a priori to estimate the value of the purely causal and purely non-causal r and s parameters. They propose a Bayesian selection procedure to select r

[4]Results with Brent are similar and available upon request.
[5]We start in 1989 for futures prices availability.

and s parameters based only on the parameters $p_{max} = r + s$. They proposed to choose p_{max} by performing lag selection in causal VAR model based on Johansen (1988) which is simply a minimisation of the information criteria. Because we will conduct a recursive forecast exercise, we have to select the value of parameters p_{max} all over our validation sample. To deals with overfitting issue in our following forecasting exercise, we select the more parsimonious information criterion which is BIC. This selection is furthermore explained by the necessity to reduce the parameters in our VAR(r,s) model which is subject to the curse of dimensionality issue. Lag selection procedure from Lanne and Luoto (2017) for RAC lead to an optimal lag p_{max} is 2.

We start by comparing the best in sample fit of possible combination for $p_{max} = 2$. Table 4.1 highlights that the better fitting model in term of likelihood from the Bayesian selection procedure of Lanne and Luoto (2017) is VAR(1,1).

Table 4.1: Loglikelihood estimation of bivariate non-causal VAR(r,s) relative to p_{max}=2 for the entire monthly dataset

VAR(1,1)	VAR(2,0)	VAR(0,2)
-325.78	-1278.36	-2583.59

Notes: VAR(r,s) refers to non-causal Bayesian vector autoregression. r denotes the causal lag parameters and s the leads parameters. The loglikelihood of the non-causal VAR(r,s) is calculated for the entire sample. Each model is estimated using 100,000 iterations, with 30,000 burns.

Table 4.2 indicates evidence in favor of fat tails, as the median posterior density value is around 4.5 degree of freedom. This suggests that the normality assumption is inappropriate thereby confirming the choice of multivariate t-distribution for the error term.

Table 4.2: Estimation of coefficients parameters of VAR(1,1) for the entire sample for Refiner Acquisition Cost of Crude oil Price (RAC)

$\Pi_{1,1}$	$\Pi_{1,2}$	$\Pi_{2,1}$	$\Pi_{2,2}$	$\Phi_{1,1}$	$\Phi_{1,2}$	$\Phi_{2,1}$	$\Phi_{2,2}$	λ
0.27	0.13	0.01	0.41	0.92	0.01	-0.03	0.95	4.22
(0.06)	(0.08)	(0.03)	(0.05)	(0.02)	(0.01)	(0.01)	(0.01)	(0.47)

Notes: RAC, respectively refers to Refiner Acquisition Cost of Crude oil Price. $\pi_{i,j}$ and $\phi_{i,j}$ $i, j \in \{1, 2\}$ respectively denotes the coefficients of Π, i.e. purely causal matrix and Φ i.e. purely non-causal matrix of our non-causal VAR(1,1). Numbers in brackets are the standard deviation of parameters estimation through our Gibbs sampler. Each model is estimated using 100,000 iterations, with 30,000 burns.

The causal terms $\Pi_{1,1}$ and $\Pi_{2,2}$ and the non-causal terms $\Phi_{1,1}$ and $\Phi_{2,2}$ are significant

(Table 4.2), they are also quite close to the unit root. These results indicate that RAC and convenience yield dynamics seems to be a non-causal AR(1,1). This provide empirically evidence of including the expectation formation mechanism in both futures ans spot prices dynamics, which is not the case in Schwartz (1997)'s models.

The anticipative cross terms (i.e $\Phi_{1,2}$ $\Phi_{2,1}$), are low but significant. This provide empirically evidences of the presence of expectations mechanism between convenience yield and RAC in oil markets. Moreover, the cross terms of the causal autoregressive parts, (i.e $\Pi_{1,2}$ $\Pi_{2,1}$) are close to zero, and insignificant (Table 4.2). These results confirm (Alquist and Kilian, 2010a; Alquist et al., 2014)'s theory, i.e. oil futures prices (also in the form of convenience yield) have a weak explanatory power on spot prices when we add an anticipative part in the specification.

4.4.3 The predictive power of the convenience yield

In this section we provide empirical evidence of predictive power of convenience yield in the oil market. We proceed by assessing the forecasting performance of our model considered above by producing a pseudo-out-of-sample forecast. Models are estimated recursively from the first observation of training sample. For RAC, the learning period is from 1989:06 to 2005:12, and the out-of-sample exercise lasts from 2006:01 to 2019:05. At each iteration, we construct the forecast density for the models as:

$$P\left(\hat{y}_{t+h} \mid y_t\right) = \int P\left(\hat{y}_{t+h} \mid y_t, \Psi_{t+h}\right) P\left(\hat{\Psi}_{t+h} \mid \Psi_t, y_t\right) P\left(\hat{\Psi}_t \mid y_t\right) \mathrm{d}\Psi \tag{4.10}$$

where $h = 1, 2, \cdots, H$ and Ψ denote the model parameters. $P\left(\hat{\Psi}_t \mid y_t\right)$ represents the posterior density of the parameters which is obtained via the MCMC simulation. $P\left(\hat{y}_{t+h} \mid y_t, \Psi_{t+h}\right)$ and $P\left(\hat{\Psi}_{t+h} \mid \Psi_t, y_t\right)$ denote the density forecast of the data and the parameters that can be obtained by simulation. The point forecast is therefore the median of the forecast density. Considering the accuracy of point forecasts, we use the *Mean Square Prediction Errors* (*MSPE*s) to compare accuracy of the forecasts to those are generated from structural models (Baumeister and Kilian, 2012, and others).

The *MSPE* for $h = 1, \cdots, H$ is computed as:

$$MSPE_h = \frac{1}{T-R} \sum_{t=R}^{T-1} \left(\hat{y}_{h,t+1|t} - y_{h,t+1} \right)^2 \tag{4.11}$$

where T is the number of observations, R is the length of the rolling window and $\hat{y}_{h,t+1|t}$ are the individual forecasts of commodity price forecasts and $y_{h,t+1}$ are average of commodities price forecasts. For this purpose we report the *MSPE* of each specification as well as the *MSPE* ratio relative to no-change forecast. Moreover, in order to statistically assess the best forecasting model, we perform both conditional and unconditional pairwise model comparisons based on the test procedure of Giacomini and White (2006). H parameters represents the horizon ahead of the forecast.

All our Bayesian model are estimated using Markov Chain Monte Carlo (MCMC) and Gibbs sampler. Our reported results are based on 100,000 Gibbs replications discarding the first 30,000 as burn-in.[6]

To prove that our bivariate framework is better than the univariate one, we also estimate different signification the univariate model. We have implemented the no-change (NC), basically assimilate to "martingale" model. We estimate, Bayesian autoregression model with a lag parameter called r, we defined the following notation for this model BAR(r). To tackle the issue of the forecast accuracy could be issued from the t-distribution, we also include the Bayesian autoregression model with t-distributed shocks, noted t-BAR(r). We also implement the univariate restriction of the previous VAR(r,s), this model is called non-causal autoregressive model, the associated notation for the rest of the paper is AR(r,s). The estimation and forecasting procedure is based on Lanne et al. (2012).

For the multivariate framework, we estimate Bayesian VAR(r). Similar to the univariate framework, we estimate t-distributed shocks Bayesian VAR, with the following notation t-VAR(r). This model is also estimated by using MCMC, Gibbs sampler and prior choice for both t-distribution and VAR parameters are based on Chiu et al. (2017a).

[6]Estimation time in high computing environment is closed to 6 hours with 80 cores.

To compare accuracy of the forecasts to those that are generated from structural models, we re-calibrate the four-variables VAR model in Baumeister and Kilian (2012) including: Real RAC, World Oil Production, Stocks and Kilian's index of economic activity.[7], with the extended dataset from June 1989 to May 2019, and estimate the associated recursive forecasts. The notation for this "baseline" model is VAR$_K$(24).

Table 4.3: Recursive MSPE ratio relative to the no-change forecast for Imported Refiner Acquisition Cost of Crude oil price (RAC)

Horizons:	1	3	6	9	12
NC (MSPE)	13.98	37.10	74.18	102.00	122.31
BAR(1)	1.00	1.00	0.99	0.98	0.98
t-BAR(1)	0.95	0.98	1.01	1.04	1.06
AR(1,1)	1.60	1.41	1.30	1.27	1.26
VAR$_{CY}$(2)	0.84*	0.95	1.03	1.07	1.09
t-VAR$_{CY}$(2)	0.84*	0.90	0.92	0.90	0.89
VAR$_{CY}$(1,1)	**0.69****	**0.82***	**0.83***	**0.81***	**0.78***
VAR$_K$(24)	0.92	0.94	0.92	1.02	1.07

Notes: The notation BAR, t-BAR, AR(r,s), VAR$_X$(r) and VAR$_X$(r,s) refers respectively to univariate Bayesian autoregression model, univariate Bayesian autoregression model with t-distributed errors, Univariate Non causal Bayesian autoregression model, Bayesian VAR, Bayesian VAR with t-distributed errors and Non-causal Bayesian vector autoregression. r denotes the causal lag parameters and s the leads parameters. CY refers to the convenience yield. VAR$_K$ refers to the four-variable VAR model in Baumeister and Kilian (2012) with extended dataset. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. baseline model is the no-change forecasts. Each model is estimated using 100,000 iterations, with 30,000 burns. ***, ** and *** indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test. The bold text in cells indicates the best model for each forecasting horizons.

Table 4.3, shows that the best models is VAR(1,1) with the CY. This results highlights the fact that our empirical specification which takes into account expectations leads to a reduction of the *MSPE*: 31% at the horizon 1 and up to 22% at horizon 12. The non-causal VARs are statistically better at all horizons than the no-change (NC). It is also better than existing models (VAR$_K$) at all horizons.

Univariate models (BAR, t-BAR, AR(1,1)) have a lower forecasting performance than multivariate models (VAR(r,s), t-VAR(r)). Adding future prices in the form of convenience yield improves the forecasts which is consistent with Alquist et al. (2014). Concerning the comparison with structural literature models. We see that the Baumeister and Kilian (2012) model (VAR$_K$) on the extended dataset has a performance similarly to the no-change (NC).

[7]Available on Kilian's website.

The variables included in this structural model are not forward-looking except the RAC. It therefore captures well the market fundamentals but can misspecified the anticipative parts in its identification. Table 4.3 also shows that taking into account the t-distribution increases forecast accuracy whether in univariate (t-AR) or multivariate (t-VAR). This is consistent with Chiu et al. (2017a), which shows that allowing the t-distribution helps to capture extreme events and part of the volatility.

Table 4.4: Recursive MSPE ratio relative to the no-change forecast for West Texas Intermediate Crude oil Spot price (WTI)

Horizons:	1	3	6	9	12
NC (MSPE)	10.54	31.62	66.88	94.76	114.83
BAR(1)	1.00	1.00	0.99	0.97	0.96
t-BAR(1)	1.60	1.37	1.25	1.23	1.24
AR(1,1)	1.62	1.38	1.26	1.23	1.24
VAR(2)	1.12	1.12	1.14	1.16	1.16
t-VAR(2)	0.86*	0.92	0.92	0.90	0.87*
VAR(1,1)	**0.65****	**0.76***	**0.77***	**0.74****	**0.73****

Notes: The notation BAR, t-AR, BAR(r,s), VAR(r) and VAR(r,s) refers respectively to univariate Bayesian autoregression model, univariate Bayesian autoregression model with t-distributed errors, Univariate non-causal Bayesian autoregression model, Bayesian VAR, Bayesian VAR with t-distributed errors and non-causal Bayesian VAR. r denotes the causal lag parameters and s the leads parameters. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The baseline model is the no-change forecasts. Each model is estimated using 100,000 iterations, with 30,000 burns. ***, ** and *** indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test. The bold text in cells indicates the best model for each forecasting horizons.

RAC is estimated in Real-Time, as in Baumeister and Kilian (2012) to improve the forecasting accurracy. By construction, RAC is subject to revision and it requires a significant amount of construction work. As robustness check, we have estimated the same econometrics specification by trying to predict the average monthly price of West Texas Intermediate (WTI) over the same period. We change our bivariate dataset which include monthly price of West Texas Intermediate (WTI) and the associated convenience yield (PCA).

One advantage is that: the WTI spot price is a financial variable, therefore by definition available in Real-Time and not very sensitive to revisions. This empirical exercise therefore refutes the hypothesis of the formation of anticipations in the nowcasting techniques and RAC revision. Moreover, future prices for WTI are available quite easily. Note that for the rest of the paper, VAR now means spot prices and the associated convenience yield.

Table 4.4 confirms our hypothesis, VAR(1,1) improves the prediction from 35% at horizon 1 to 27% at horizon 12 against the no-change forecast (NC) associated with WTI monthly average spot price.

4.4.4 An extension to the U.S natural gas markets

The use of convenience yield in Schwartz (1997) is build on the *the storage theory*, as stated in Alquist et al. (2014) and Working (1949). This theory shows that convenience yield is consistent with the price to pay for storing a barrel of oil during a fixed period. This is based on the assumption of a liquid and global market with low transportation costs. In this section we test the robustness of our approach the U.S Natural Gas market, which is a more local market with higher transportation costs (Rosendahl and Sagen, 2009). We use monthly average Henry Hub spot prices and along with the associated futures prices to construct the convenience yield as for oil. Estimation procedure for the VAR(r,s) is the same as above. Data sample start in 1997:01 and ends in 2019:05, as the Henry Hub futures prices start to be available only in 1997. The optimal lag estimated by the Bayesian lag selection procedure from Lanne and Luoto (2017), is $p_{max} = 2$, and the best in sample fit is for a VAR(1,1) model, which is similar to the results on RAC.

Table 4.5: Caption Estimation of coefficients parameters of VAR(1,1) for the entire sample for Henry Hub Natural Gas Spot Price (HH)

$\Pi_{1,1}$	$\Pi_{1,2}$	$\Pi_{2,1}$	$\Pi_{2,2}$	$\Phi_{1,1}$	$\Phi_{1,2}$	$\Phi_{2,1}$	$\Phi_{2,2}$	λ
0.29	0.17	-0.21	0.02	0.54	-0.06	0.14	0.97	2.75
(0.05)	(0.09)	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)	(0.02)	(0.40)

Notes: $\pi_{i,j}$ and $\phi_{i,j}$ $i,j \in \{1,2\}$ respectively denotes the coefficients of Π, i.e. purely causal matrix and Φ i.e. purely non-causal matrix of our non-causal VAR(1,1). Numbers in brackets are the standard deviation of parameters estimation through our Gibbs sampler. Each model is estimated using 100,000 iterations, with 30,000 burns.

Table 4.5 indicates strong evidence in favor of fat tails as the estimated posterior median of DOF parameters λ is equal to 2.75 degree of freedom with an associated standard deviation around 0.40. This confirms the presence of non-Gaussianity in the error terms. Estimation of λ for gas is lower than the estimated degree-of-freedom of RAC (around 4.5). It seems that gas prices are more sensitive to extreme values than oil. This fact can be explained by several factors related to gas markets. First, both natural gas demand and prices are highly dependant to extreme weather events (temperature).

Table 4.5 shows that: Estimation of the causal, i.e Π, and non-causal, i.e, Φ, parts are all significant except for the purely autoregressive component related to the convenience yield, i.e, $\Pi_{2.2}$. This confirms the presence of expectations on the dynamics of futures and spot prices for both commodity (See Table 4.2 for RAC)

Table 4.5, also shows that, like for oil, cross terms are also significant. More precisely, $\Phi_{1.2}$ and $\Phi_{2.1}$ have a larger estimated amplitude than oil, respectively -0.06 and 0.14 for natural gas and 0.03 and -0.03. These empirical results can be related to transportation/storage costs which are higher in the gas market than in the oil market (Rosendahl and Sagen, 2009). U.S natural gas stocks adjustment, in response to demand and/or supply shocks which are reflected in the spot price is less flexible, take more time and much more expensive than in the case of oil. Expectations formed by agents, about the value to hold a "barrel" of natural gas is more important than the rapidity of convenience yield price adjustment.

We test whether future prices using as convenience yield have a predictive power in the case of natural gas. Our forecasting procedure is similar to the previous one for the RAC but due to data available issue, the learning period is from 1997:01 to 2011:12, and the out-of-sample is from 2012:01 to 2019:05. For structural models, we estimate a four-variable VAR model including: the U.S. real price of natural gas, the U.S. natural gas production, the U.S industrial production and the real price of crude oil. We also consider our newly constructed temperature index as an additional exogenous variable from Chapter 2 noted VARX.

Table 4.6,confirms the predictive power of the convenience yield attached to the spot price. Our empirical specification, allows to improve from 21% to 32% whatever the horizon, the forecast quality compared to the no-change forecast (NC). Adding t-distribution helps in multivariate (t-VAR) with an average reduction in the MSPE of about 15% across horizons, compared with the no-change forecasts (NC).

Table 4.6, shows that our structural models VARX $MSPE$ are lower than the $MSPE$ of the VAR(1,1) at long-term horizon. The $MSPE$ reduction is high as 7% at 6 months-ahead

Table 4.6: Recursive MSPE ratio relative to the no-change forecast for Henry Hub Natural Gas Spot Price (HH)

Horizons:	1	3	6	9	12
NC (MSPE)	13.40	22.35	34.51	47.05	58.39
BAR(1)	0.98	0.97	0.97	0.98	1.01
t-BAR(1)	1.43	1.27	1.20	1.18	1.18
AR(1,1)	1.44	1.27	1.21	1.19	1.18
VAR(2)	0.88*	0.98	1.07	1.11	1.12
t-VAR(2)	0.79**	0.85*	0.85*	0.86*	0.87*
VAR(1,1)	**0.68****	0.76**	0.73**	0.74**	0.79**
VARX(12)	1.00	0.84**	**0.66*****	**0.55*****	**0.49*****

Notes: The notation BAR, t-BAR, AR(r,s), VAR(r) and VAR(r,s) refers respectively to univariate Bayesian autoregression model, univariate Bayesian autoregression model with t-distributed errors, Univariate non-causal Bayesian autoregression model, Bayesian VAR, Bayesian VAR with t-distributed errors and non-causal Bayesian VAR. r denotes the causal lag parameters and s the leads parameters. VARX refers to the four-variable VAR model including: the U.S. real price of natural gas, the U.S. natural gas production, the U.S industrial production and the real price of crude oil with our newly constructed temperature index as an additional exogenous variable from Chapter 2. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The baseline model is the no-change forecasts. Each model is estimated using 100,000 iterations, with 30,000 burns. ***, ** and *** indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test. The bold text in cells indicates the best model for each forecasting horizons.

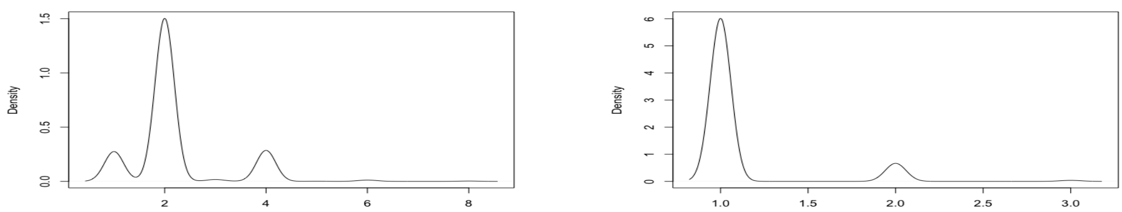
and 30% 12 months-ahead. In short-term, the VAR(1,1) have better forecasting accuracy by 32% (1 months-ahead) and 8% (3 months-ahead). This is explain by the fact that in short term the convenience yield value, is closer to the real cost to adjust the stock than in long-term where uncertainty is much higher.

4.4.5 Does the performance hold for higher frequencies?

Optimal lag determination

In the previous section, we see that the expectations related to the convenience yield are formed in a similar way for oil and gas. Nevertheless, executing a pseudo-exercise forecasts with weekly and daily data leads to technical difficulties notably the time of the Bayesian estimation became much longer. We adapted the previous estimation and forecast procedures to tackle estimation time. We choose to use a rolling window to reduce the dataset. To calibrate the size of the window, we estimate the optimal number $p_{max} = (r + s)$ of lags based on the procedure of Johansen (1988) for different window sizes with a fixed step. We looked at the average of the p_{max} as a function of the chosen estimation window over our entire data set. Based on this, we looked at the stability for each window of the p_{max} obtained. For the sake of clarity we explain the procedure only for WTI. For weekly data p_{max} is equal to 2 or 4 depending on the size of the window. We have kept the window that maximizes in probability $p_{max} = 2$ across all sliding samples (Figure 4.1, Panel A), the size of the optimal window is 250 (approximately 5 years of quotations). For daily data, p_{max} is most often equal to 1, which is not surprising since the AR(1) model is a discrete time solution of continuous time processes of the Ornstein-Uhlenbeck Schwartz (1997) 's type.[8] To find the optimal windows we have maximized the probablility of $p_{max} = 2$ (Figure 4.1, Panel B), and find an optimal window size of 750 (approximately 3 years of quotations).

Figure 4.1: Estimated density of p_{max} for West Texas Intermediaire (WTI) with optimal rolling windows



Panel A: Weekly optimal window=250 Panel B: Daily optimal window=750

Results

In the Table 4.7. As a reminder, we have added the estimated coefficients on the monthly data. Overall the DOF (λ) is always significant whatever the frequency of the data and around 4 for WTI and 2 for HH. This confirms our choice of t-distribution even at higher frequencies.

For WTI, Table 4.7 shows the coefficients are all significant except $\pi_{2.1}$ for weekly data. It is important to stress that the coefficients $\pi_{1.1}$, $\pi_{2.2}$, $\phi_{1.1}$ and $\phi_{2.2}$ remain stable between the monthly and weekly data, however the cross terms $\phi_{1.2}$ and $\phi_{2.1}$ are larger in weekly than in monthly data, which shows that expectations are stronger at shorter horizons. For the daily data, we see that the $\pi_{1.1}$, $\pi_{2.2}$, $\phi_{1.1}$ and $\phi_{2.2}$ are very close to 1 which is by definition a unit

[8]Which is one of the most used for commodity option pricing

Table 4.7: Estimation of coefficients parameters of VAR(1,1) for the entire sample for Henry Hub Natural Gas Spot Price (HH) and West Texas Intermediaire (WTI)-Monthly, Weekly and Daily

	$\Pi_{1,1}$	$\Pi_{1,2}$	$\Pi_{2,1}$	$\Pi_{2,2}$	$\Phi_{1,1}$	$\Phi_{1,2}$	$\Phi_{2,1}$	$\Phi_{2,2}$	λ
WTI									
Monthly	0.29	-0.02	-0.39	0.51	0.23	0.07	0.45	0.90	4.33
	(0.05)	(0.10)	(0.03)	(0.06)	(0.11)	(0.02)	(0.10)	(0.02)	(0.76)
Weekly	0.25	-0.14	-0.04	0.40	0.65	0.18	-0.48	0.81	3.61
	(0.05)	(0.05)	(0.10)	(0.08)	(0.06)	(0.05)	(0.23)	(0.13)	(0.62)
Daily	0.99	0.01	-0.02	0.97	0.83	-0.01	0.02	0.83	4.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)	(0.04)	(0.45)
HH									
Monthly	0.29	0.17	-0.21	0.02	0.54	-0.06	0.14	0.97	2.75
	(0.05)	(0.09)	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)	(0.02)	(0.40)
Weekly	0.01	0.01	-0.01	0.17	0.96	0.04	0.01	0.96	2.06
	(0.04)	(0.07)	(0.03)	(0.05)	(0.02)	(0.02)	(0.01)	(0.01)	(0.06)
Daily	-0.03	-0.05	-0.01	-0.06	-0.96	-0.01	-0.04	0.98	2.15
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.29)

Notes: $\pi_{i,j}$ and $\phi_{i,j}$ $i,j \in \{1,2\}$ respectively denotes the coefficients of Π, i.e. purely causal matrix and Φ i.e. purely non-causal matrix of our non-causal VAR(1,1). Numbers in brackets are the standard deviation of parameters estimation through our Gibbs sampler. Each model is estimated using 100,000 iterations, with 30,000 burns.

root inside the circle. This results leads to the presence of a purely anticipative martingale process.

For natural gas and with weekly data, Table 4.7 shows the coefficients $\pi_{2.2}$, $\phi_{1.1}$, $\phi_{1.2}$, $\phi_{2.2}$ are significant. The causal autoregressive part is non-significant and close to zeros. It seems that gas prices have a significant forward-looking behavior. The coefficients $\phi_{1.1}$ and $\phi_{2.2}$ are very close to 1 and to the anticipative unit root. According to Table 4.7 the results are robust to the passage to the daily frequency.

We are now going to investigate if whether the VAR(1,1) forecasting performance also exist to higher data frequency (week and days). We have changed our forecast horizon to 4 for the week and 7 for the daily data.

Our approach is robust to a change in frequency, Table 4.8 and 4.9 confirms the good performance of the VAR(1,1).

For oil, Table 4.8 highlights some other complementary results: we find that BAR(1) and t-BAR(1), *MSPE* is lower than the *MSPE* of no-change for long-term horizon (>1 for both

Table 4.8: Recursive MSPE ratio relative to the no-change forecast for West Texas Intermediate Crude oil Spot price (WTI) with Weekly and Daily data

Weekly							
Horizons:	1	2	3	4			
NC (MSPE)	0.09	0.18	0.27	0.36			
BAR(1)	0.93	0.60***	0.44***	0.35***			
t-BAR(1)	1.05	0.64***	0.46***	0.36***			
AR(1,1)	16.39	13.06	11.19	9.89			
VAR(2)	8.34	5.95	4.89	4.23			
t-VAR(2)	1.15	1.15	1.16	1.18			
VAR(1,1)	0.80**	0.72**	0.68**	0.65**			
Daily							
Horizons:	1	2	3	4	5	6	7
NC (MSPE)	0.06	0.08	0.10	0.11	0.13	0.15	0.16
BAR(1)	0.98	0.72**	0.56***	0.46***	0.39***	0.34***	0.30***
BAR(1)	0.98	0.72***	0.56***	0.46***	0.39***	0.34***	0.29***
AR(1,1)	6.76	5.29	4.56	4.10	3.76	3.51	3.31
VAR(2)	13.49	10.38	8.80	7.80	7.07	6.52	6.08
t-VAR(2)	1.01	1.01	1.01	1.02	1.02	1.03	1.03
VAR(1,1)	0.97	0.91*	0.87**	0.83**	0.81**	0.79**	0.77**

Notes: The notation BAR, t-BAR, AR(r,s), VAR(r) and VAR(r,s) refers respectively to univariate Bayesian autoregression model, univariate Bayesian autoregression model with t-distributed errors, Univariate Non-causal Bayesian autoregression model, Bayesian VAR, Bayesian VAR with t-distributed errors and Non-causal Bayesian VAR. r denotes the causal lag parameters and s the leads parameters. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The baseline model is the no-change forecasts. Each model is estimated using 100,000 iterations, with 30,000 burns. ***, ** and *** indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test.

weekly and daily data). Univariate models seems to have a better forecasting performance than any of multivariate specifications, except VAR(1,1). In multivariate case, fat tails, surprisingly, brings the quality of forecasting to a similar level than the no-change. Table 4.9, confirms the results for natural gas. Only major difference for natural gas is the poor accuracy of VAR(1,1). Moreover, it is important to stress, that in Table 4.8, $MSPE$ at horizon 5 using daily data (0.81), is equal to the $MSPE$ at horizon 1 for weekly data (0.80). This shows the robustness of our approach.

In conclusion, despite the parsimony of our model, the quality of the forecasts remains interesting.

Table 4.9: Recursive MSPE ratio relative to the no-change forecast for Henry Hub Natural Gas Spot Price (HH) with Weekly and Daily data

			Weekly				
Horizons:	1	2	3	4			
NC (MSPE)	0.14	0.19	0.23	0.26			
BAR(1)	1.28	1.00	0.86**	0.84**			
t-BAR(1)	24.21	14.46	10.62	8.59			
AR(1,1)	11.43	8.22	6.39	5.36			
VAR(2)	1.80	1.48	1.31	1.23			
t-VAR(2)	1.05	1.06	1.04	1.04			
VAR(1,1)	0.80**	0.72**	0.68**	0.65**			
			Daily				
Horizons:	1	2	3	4	5	6	7
NC (MSPE)	0.06	0.08	0.10	0.11	0.13	0.15	0.16
BAR(1)	2.01	1.34	0.99	0.78**	0.64***	0.53***	0.45***
t-BAR(1)	1.88	1.28	0.96	0.76**	0.62***	0.52***	0.45***
AR(1,1)	9.24	7.05	5.92	5.18	4.64	4.21	3.87
VAR(2)	15.20	8.62	6.02	4.58	3.65	3.00	2.53
t-VAR(2)	23.53	16.87	13.80	11.87	10.49	9.44	8.62
VAR(1,1)	2.24	2.17	2.15	2.10	2.06	2.01	1.97

Notes: The notation BAR, t-BAR, AR(r,s), VAR(r) and VAR(r,s) refers respectively to univariate Bayesian autoregression model, univariate Bayesian autoregression model with t-distributed errors, Univariate Non-causal Bayesian autoregression model, Bayesian VAR, Causal Bayesian VAR with t-distributed errors and Non-causal Bayesian VAR. r denotes the causal lag parameters and s the leads parameters. Entries less than 1 indicate that forecasts from the indicated model are more accurate than forecasts from the associated baseline model. The baseline model is the no-change forecasts. Each model is estimated using 100,000 iterations, with 30,000 burns. ***, ** and *** indicate MSPE ratios are significantly different from 1 at 1%, 5% and 10%, according to both conditional and unconditional Giacomini and White (2006) test.

4.5 Conclusion

This paper presents empirical results confirming, the predictive power of convenience yield. Our main contribution, is to use bivariate Bayesian Non-causal VAR framework from Lanne and Luoto (2016), to estimate the expectations mechanism underlined in the theory of storage. With this new features, we provide empirical evidence that real-time forecasts of real oil prices can be significantly more accurate than the no-change forecast and remarkably more accurate than real-time forecasts generated by existing structural models relying on Bayesian VAR such as those in (Baumeister and Kilian, 2012, among others).

We provide evidence that this methodology could be extend to different commodity such as natural gas. We use the Henry Hub Natural gas spot prices. We also investigate the robustness of our approach to weekly and daily data. There are several promising avenues of research which can improve this preliminary version of the paper. The ongoing work in this sense is to provides some theoretical demonstration to prove the existence of a Schwartz (1997)'s 2-factors models equation solution that include expectations. We investigate, the fact that our bivariate non-causal VAR(1,1) could be a discrete time approximation of this solution. Notwithstanding the value of our findings, it should be borne in mind that our empirical analysis can be extended for others purposes: valuation of option, portefolio/risk management.

General conclusion

The contribution of this thesis is twofold. First, we provide a quantitative analysis of the economic determinants of energy demand. Second, we develop demand forecasting models. 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 balancing purposes, and (iii) the possibility to use existing gas infrastructures to supply short-term, flexibility services to a renewable-dominated power sector.

The classic identification problem stems from the fact that demand and price adjust simultaneously. In the literature, practitioners often focus on the supply side. Chapter 1 provides evidence that the information in these prices is substantial enough to provide reasonably accurate predictions of the next day's consumption of natural gas and at least much better than those of TSO's. 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. 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 and apply it to the case of France. Our results have shown that its forecasting performance outperforms the tools routinely used by infrastructure operators. Our results suggest that accounting for the information contained in day-ahead prices represents a promising avenue to improve the performance of these demand forecasts. 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 relationship 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. One promising avenue of this research is to deals with the other classical demand modelisation: aggregation (See).

In chapter 2, we estimate various Bayesian VAR models following the seminal contribution by Baumeister and Kilian (2012) and more recently Baumeister et al. (forthcoming) with an emphasize on the demand side which is shown to be highly relevant in the case of gas. Moreover, the use of temperature data that is made possible by the regional feature of the U.S. gas market permits to further include a true real-time proxy for demand in our econometric specifications. Our results, that are the first of this kind for natural gas and are likely to serve as future benchmarks for economic research in the field of energy economics, point to large improvements relative to the no-change forecast. We provide evidence that models that comprise temperature and allow for stochastic volatility and fat tails deliver the best forecasts. The role of temperature, in particular, is a novel finding that deserves to be further investigated. Another strand of the literature initiated in Alquist and Kilian (2010b) deals with the information content of futures prices to make predictions about commodity prices. While we do not elaborate on this possibility for natural gas in the present paper, a related approach in Thomas (2020) is noteworthy. More specifically, Thomas (2020) relies on a non causal bivariate VAR for crude oil and natural gas. In his model, the two variables under consideration are the energy price and the convenience yield. The latter variable is included on the economic ground that it does proxy for expectations in the derivatives market. Surprisingly, such a very parsimonious model exhibits highly accurate forecasts at horizons up to several months.

Chapter 3, we propose a methodology to identify news shock under nonfundamentalness in oil markets. We focus on oil market to be able to compare our results to the previous literature, but this approach is easily applicable to the natural gas markets, with the dataset from Chapter 2. Indeed, purely autoregressive VAR models have become prominent in empirical works on global oil market in spite of the presence of nonfundamentalness problem which may arise as result of the informational insufficiency in the small-scale VAR model. Until now, nonfundamentalness issue on the global oil market has been addressed by either augmenting small-scale VAR models by additional variables or latent factors, or using external instrument or proxies leading to more credible identification scheme. In this paper we dealt with this issue by employing the non-causal VAR model on standard global oil market, namely global oil production, global economic activity, oil stocks and real oil price and US macroeconomic variables in order to analyse the effect of the oil supply news shock. We showed, first, that the nonfundamental representation is supported by the data, justifying therefore the use of Non-Causal VAR as an option dealing with the information deficiencies when modelling global oil market. Second, follow Nelimarkka (2017a) we identified an oil supply news shock as a shock that drives global oil production the most for a finite time horizon. We further showed that our identified oil supply news shock is anticipated by forward-looking variables before the shock materialises, highlighting the prominent role of expectations in propagating the shock. We documented also that a negative oil supply news shock results in abrupt and permanent reaction in global oil production, global economic activity and in oil inventory. However, the oil supply shock has only a limited effect on oil price. Moreover, news shock about oil supply shortfalls do have macroeconomic consequences as it causes a substantial decline in US industrial production. Finally, evidence on the prominent role of the expectation channel is confirmed by the reaction of consumer and business confidence indicators over the anticipatory lags before the shock materialises. While NC-VAR represents a promising approach dealing with nonfundamentalness problem, the issue of identification of structural shocks remains challenging task. The lack of development of structural NC-VAR is unfortunate, preventing the analysis of the effect of further interesting shocks using credible identification scheme. As further research, it would be worthwhile to explore the impact of the oil demand shock on global oil market and macroeconomic variables.

Chapter 4 presents empirical results confirming, the predictive power of convenience yield.

Our main contribution, is to use bivariate Bayesian Non-causal VAR framework from Lanne and Luoto (2016), to estimate the expectations mechanism underlined in the theory of storage. With this new features, we provide empirical evidence that real-time forecasts of real oil prices can be significantly more accurate than the no-change forecast and remarkably more accurate than real-time forecasts generated by existing structural models relying on Bayesian VAR such as those in (Baumeister and Kilian, 2012, among others). We provide evidence that this methodology could be extend to different commodity such as natural gas. We use the Henry Hub Natural gas spot prices. We also investigate the robustness of our approach to weekly and daily data. There are several promising avenues of research which can improve this preliminary version of the paper. The ongoing work in this sense is to provides some theoretical demonstration to prove the existence of a Schwartz (1997)'s 2-factors models equation solution that include expectations. We investigate, the fact that our bivariate non-causal VAR(1,1) could be a discrete time approximation of this solution. Notwithstanding the value of our findings, it should be borne in mind that our empirical analysis can be extended for others purposes: valuation of option, portefolio/risk management.

Bibliography

Abada, I., Massol, O., 2011. Security of supply and retail competition in the european gas market: some model-based insights. Energy Policy 39, 4077–4088.

ACER-ENTSOG, 2014. ACER-ENTSOG Report on the early implementation of the Balancing Network Code. Technical Report. European network of transmission system operators for gas: Brussels.

Adeyemi, O.I., Hunt, L.C., 2014. Accounting for asymmetric price responses and underlying energy demand trends in oecd industrial energy demand. Energy Economics 45, 435 – 444.

Agnolucci, P., De Lipsis, V., Arvanitopoulos, T., 2017. Modelling UK. sub-sector industrial energy demand. Energy Economics 67, 366–374.

Alessi, L., Barigozzi, M., Capasso, M., 2011. Non-fundamentalness in structural econometric models: A review. International Statistical Review 79, 16–47.

Alquist, R., Bauer, G., de los Rios, A.D., 2014. What does the convenience yield curve tell us about the crude oil market? Staff Working Papers 14-42. Bank of Canada.

Alquist, R., Kilian, L., 2010a. What do we learn from the price of crude oil futures? Journal of Applied Econometrics 25, 539–573.

Alquist, R., Kilian, L., 2010b. What do we learn from the price of crude oil futures? Journal of Applied Econometrics 25, 539–573.

Alquist, R., Kilian, L., Vigfusson, R.J., 2013. Chapter 8 - forecasting the price of oil, in: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Elsevier. volume 2 of *Handbook of Economic Forecasting*, pp. 427 – 507.

Andrews, D., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.

Ang, B., 1995. Decomposition methodology in industrial energy demand analysis. Energy 20, 1081 – 1095.

Antolín-Díaz, J., Rubio-Ramírez, J.F., 2018. Narrative sign restrictions for svars. American Economic Review 108, 2802–29.

Aras, H., Aras, N., 2004. Forecasting residential natural gas demand. Energy Sources 26, 463–472.

Arezki, R., Ramey, V.A., Sheng, L., 2017. News Shocks in Open Economies: Evidence from Giant Oil Discoveries. The Quarterly Journal of Economics 132, 103–155.

Arvesen, o., Medbø, V., Fleten, S.E., Tomasgard, A., Westgaard, S., 2013. Linepack storage valuation under price uncertainty. Energy 52, 155–164.

Bachmeier, L.J., Griffin, J.M., 2006. Testing for market integration crude oil, coal, and natural gas. Energy Journal 27, 55–71.

Bagnai, A., Ospina, C.A.M., 2018. Asymmetries, outliers and structural stability in the US gasoline market. Energy Economics 69, 250–260.

Barndorff-Nielsen, O.E., Shephard, N., 2001. Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics. Journal of the Royal Statistical Society. Series B (Statistical Methodology) 63, 167–241.

Barsky, R.B., Sims, E.R., 2011. News shocks and business cycles. Journal of Monetary Economics 58, 273–289.

Baumeister, C., Guérin, P., Kilian, L., 2015. Do high-frequency financial data help forecast oil prices? the midas touch at work. International Journal of Forecasting 31, 238 – 252.

Baumeister, C., Hamilton, J.D., 2019a. Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. American Economic Review 109, 1873–1910.

Baumeister, C., Hamilton, J.D., 2019b. Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. American Economic Review 109.

Baumeister, C., Kilian, L., 2012. Real-time forecasts of the real price of oil. Journal of Business & Economic Statistics 30, 326–336.

Baumeister, C., Kilian, L., 2014. What central bankers need to know about forecasting oil prices. International Economic Review 55, 869–889.

Baumeister, C., Kilian, L., 2015. Forecasting the real price of oil in a changing world: A forecast combination approach. Journal of Business & Economic Statistics 33, 338–351.

Baumeister, C., Kilian, L., 2016. Understanding the Decline in the Price of Oil since June 2014. Journal of the Association of Environmental and Resource Economists 3, 131–158.

Baumeister, C., Kilian, L., Lee, T., 2017a. Inside the crystal ball: New approaches to predicting the gasoline price at the pump. Journal of Applied Econometrics 32, 275–295.

Baumeister, C., Kilian, L., Lee, T.K., 2014. Are there gains from pooling real-time oil price forecasts? Energy Economics 46, S33–S43.

Baumeister, C., Kilian, L., Lee, T.K., 2017b. Inside the crystal ball: New approaches to predicting the gasoline price at the pump. Journal of Applied Econometrics 32, 275–295.

Baumeister, C., Kilian, L., Zhou, X., 2018. Are product spreads useful for forecasting oil prices? an empirical evaluation of the verleger hypothesis. Macroeconomic Dynamics 22, 562–580.

Baumeister, C., Korobilis, D., Lee, T., forthcoming. Energy markets and global economic conditions. Review of Economics and Statistics .

Baumeister, C., Korobilis, D., Lee, T.K., 2020. Energy Markets and Global Economic Conditions. Working Paper 27001. National Bureau of Economic Research.

Baumeister, C., Peersman, G., 2013. Time-varying effects of oil supply shocks on the us economy. American Economic Journal: Macroeconomics 5, 1–28.

Bańbura, M., Giannone, D., Reichlin, L., 2010. Large Bayesian vector auto regressions. Journal of Applied Econometrics 25, 71–92.

Beaudry, P., Portier, F., 2006. Stock prices, news, and economic fluctuations. American Economic Review 96, 1293–1307.

Beaudry, P., Portier, F., 2014. News-driven business cycles: Insights and challenges. Journal of Economic Literature 52, 993–1074.

Behme, A.D., 2011. Distributional properties of solutions of dv t = v t-du t dl t with lévy noise. Advances in Applied Probability 43, 688–711.

Bello, A., Bunn, D.W., Reneses, J., noz, A.M., 2017. Medium-term probabilistic forecasting of electricity prices: A hybrid approach. IEEE Transactions on Power Systems 32, 334–343.

Berrisford, H.G., 1965. The relation between gas demand and temperature: A study in statistical demand forecasting. Journal of the Operational Research Society 16, 229–246.

Bessembinder, H., Chan, K., 1992. Time-varying risk premia and forecastable returns in futures markets. Journal of Financial Economics 32, 169 – 193.

Bohi, D., 1981. Analyzing demand behavior: A study of energy elasticities. John Hopkins University Press.

Bordignon, S., Bunn, D.W., Lisi, F., Nan, F., 2013. Combining day-ahead forecasts for british electricity prices. Energy Economics 35, 88 – 103.

Brennan, M.J., Hughes, P.J., 1991. Stock prices and the supply of information. The Journal of Finance 46, 1665–1691.

Brown, S.P., Yücel, M.K., 2008. Deliverability and regional pricing in U.S. natural gas markets. Energy Economics 30, 2441–2453.

Bunn, D., Chevallier, J., Le Pen, Y., Sévi, B., 2017. Fundamental and financial influences on the co-movement of oil and gas prices. Energy Journal 38, 201–228.

Bunn, D.W., Gianfreda, A., 2010. Integration and shock transmissions across european electricity forward markets. Energy Economics 32, 278 – 291.

Caldara, D., Cavallo, M., Iacoviello, M., 2019. Oil price elasticities and oil price fluctuations. Journal of Monetary Economics 103, 1 – 20.

Canyurt, O.E., Ozturk, H.K., 2008. Application of genetic algorithm (GA) technique on demand estimation of fossil fuels in Turkey. Energy Policy 36, 2562–2569.

Caporin, M., Fontini, F., 2017. The long-run oil–natural gas price relationship and the shale gas revolution. Energy Economics 64, 511 – 519.

Cappé, O., Douc, R., Guillin, A., Marin, J.M., Robert, C.P., 2008. Adaptive importance sampling in general mixture classes. Statistics and Computing 18, 447–459.

CEDIGAZ, 2019. Natural Gas In The World. Annual Report. CEDIGAZ.

Chan, J.C., Eisenstat, E., 2017. Efficient estimation of bayesian varmas with time-varying coefficients. Journal of Applied Econometrics 32, 1277–1297.

Charlot, P., Darné, O., Moussa, Z., 2016. Commodity returns co-movements: Fundamentals or "style" effect? Journal of International Money and Finance 68, 130–160.

Chen, Y.C., Rogoff, K.S., Rossi, B., 2010. Can Exchange Rates Forecast Commodity Prices?*. The Quarterly Journal of Economics 125, 1145–1194.

Chevallier, J., Goutte, S., 2017. Estimation of Lévy-driven Ornstein–Uhlenbeck processes: application to modeling of CO_2 and fuel-switching. Annals of Operations Research 255, 169–197.

Chib, S., Ramamurthy, S., 2010. Tailored randomized block MCMC methods with application to DSGE models. Journal of Econometrics 155, 19–38.

Chiu, C.W.J., Mumtaz, H., Pinter, G., 2017a. Forecasting with VAR models: Fat tails and stochastic volatility. International Journal of Forecasting 33, 1124–1143.

Chiu, C.W.J., Mumtaz, H., Pintér, G., 2017b. Forecasting with VAR models: Fat tails and stochastic volatility. International Journal of Forecasting 33, 1124–1143.

Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. Journal of Finance 46, 209–237.

Cogley, T., Sargent, T.J., 2005. Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. Review of Economic Dynamics 8, 262–302.

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. Technical Report. Commission de Régulation de l'Énergie: Paris.

Crompton, P., Wu, Y., 2005. Energy consumption in China: past trends and future directions. Energy Economics 27, 195–208.

Davis, R.A., Song, L., 2020. Noncausal vector ar processes with application to economic time series. Journal of Econometrics 216, 246 – 267.

Dergiades, T., Tsoulfidis, L., 2008. Estimating residential demand for electricity in the United States, 1965–2006. Energy Economics 30, 2722 – 2730.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

Ediger, V., Akar, S., Ugurlu, B., 2006. Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model. Energy Policy 34, 3836–3846.

Ediger, V.S., Akar, S., 2007. ARIMA forecasting of primary energy demand by fuel in Turkey. Energy Policy 35, 1701–1708.

Elkhafif, M.A.T., 1996. An iterative approach for weather-correcting energy consumption data. Energy Economics 18, 221–230.

Eltony, M.N., 1996. Demand for natural gas in Kuwait: An empirical analysis using two econometric models. International Journal of Energy Research 20, 957–963.

Engle, R., Granger, C., 1987. Co-integration and error correction: Representation, estimation, and testing. Econometrica 55, 251–76.

ENTSOG, 2017. Balancing network code. Implementation and effect monitoring report. Technical Report. European Network of Transmission System Operators for Gas: Brussels.

Erdogdu, E., 2009. Natural gas demand in Turkey. Technical Report.

Farahbakhsh, H., Ugursal, V.I., Fung, A.S., 1998. A residential end-use energy consumption model for canada. International Journal of Energy Research 22, 1133–1143.

Fattouh, B., Kilian, L., Mahadeva, L., 2012. The Role of Speculation in Oil Markets: What Have We Learned So Far? CEPR Discussion Papers. C.E.P.R. Discussion Papers.

Ferrari, D., Ravazzolo, F., Vespignani, J., 2019. Forecasting energy commodity prices: A large global dataset sparse approach. CAMP Working Paper Series No 11/2019.

Forbes, K.F., Zampelli, E.M., 2014. Do day-ahead electricity prices reflect economic fundamentals? Evidence from the California ISO. Energy Journal 35, 129–144.

Forni, M., Gambetti, L., Sala, L., 2014. No news in business cycles. The Economic Journal 124, 1168–1191.

Francis, N., Owyang, M.T., Roush, J.E., DiCecio, R., 2014. A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks. The Review of Economics and Statistics 96, 638–647.

Fraumeni, B., Jorgenson, D., 1981. Relative prices and technical change. Modeling and Measuring Natural Resource Substitution, M.I.T. Press, Cambridge. pp. 17–47.

Fries, S., 2018. Anticipative alpha-stable linear processes for time series analysis : conditional dynamics and estimation. Ph.D. thesis. URL: `http://www.theses.fr/2018SACLG005/document`.

Gao, S., Hou, C., Nguyen, B., 2020. Forecasting natural gas prices using highly flexible time-varying parameter models. Tasmanian School of Business and Economics, University of Tasmania .

Geweke, J., 1978. Temporal aggregation in the multiple regression model. Econometrica 46, 643–661.

Geweke, J., 1993. Bayesian treatment of the independent student-t linear model. Journal of Applied Econometrics 8, S19–S40.

Giacomini, R., White, H., 2006. Tests of conditional predictive ability. Econometrica 74, 1545–1578.

Gianfreda, A., Bunn, D., 2018. A stochastic latent moment model for electricity price formation. Operations Research 66, 1189–1203.

Gibson, R., Schwartz, E.S., 1990. Stochastic convenience yield and the pricing of oil contingent claims. Journal of Finance 45, 959–76.

Gilchrist, S., Zakrajšek, E., 2012. Credit spreads and business cycle fluctuations. American Economic Review 102, 1692–1720.

Giulietti, M., Grossi, L., Waterson, M., 2012. A rough analysis: valuing gas storage. Energy Journal 33, 119–141.

Gopalakrishnan, A., Biegler, L.T., 2013. Economic nonlinear model predictive control for periodic optimal operation of gas pipeline networks. Computers Chemical Engineering 52, 90 – 99.

Gorucu, F.B., 2004. Evaluation and forecasting of gas consumption by statistical analysis. Energy Sources 26, 267–276.

Gospodinov, N., Ng, S., 2013. Commodity Prices, Convenience Yields, and Inflation. The Review of Economics and Statistics 95, 206–219.

Gourieroux, C., Hencic, A., Jasiak, J., 2019. Forecast Performance and Bubble Analysis in Noncausal MAR(1,1) Processes. Technical Report.

Gourieroux, C., Jasiak, J., 2017a. Noncausal vector autoregressive process: Representation, identification and semi-parametric estimation. Journal of Econometrics 200, 118 – 134.

Gourieroux, C., Jasiak, J., 2017b. Noncausal vector autoregressive process: Representation, identification and semi-parametric estimation. Journal of Econometrics 200, 118 – 134.

Gouriéroux, C., Lu, Y., 2019. Non-causal Affine Processes with Applications to Derivative Pricing. Working Papers. Center for Research in Economics and Statistics.

Green, R., Vasilakos, N., 2010. Market behaviour with large amounts of intermittent generation. Energy Policy 38, 3211 – 3220.

Greenwood-Nimmo, M., Shin, Y., Van Treeck, T., Konjunkturforschung, D., 2011. The asymmetric ARDL model with multiple unknown threshold decompositions: An application to the Phillips curve in Canada, in: The Leeds University Business School Working Paper Series.

GRID, N., 2018. National Grid Gas (NTS) System Operator Incentives. 2018/19 Supporting Information. Technical Report. National Grid UK.

Gürkaynak, R.S., Sack, B., Swanson, E., 2005. The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. American Economic Review 95, 425–436.

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

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.

Hartman, R.S., 1979. Frontiers in energy demand modeling. Annual Review of Energy 4, 433–466.

Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting 13, 281 – 291.

Hausman, C., Kellogg, R., 2015. Welfare and distributional implications of shale gas. Brookings Papers on Economic Activity 46, 71–139.

Heather, P., Petrovich, B., 2017. European traded gas hubs: An updated analysis on liquidity, maturity and barriers to market integration. Technical Report. Oxford Institute for Energy Studies.

Helm, D., 2012. The Kyoto approach has failed. Nature 491, 663–665.

Hendry, D.F., Juselius, K., 2000. Explaining cointegration analysis: Part 1. Energy Journal 21, 1–42.

Heyde, C.C., Leonenko, N.N., 2005. Student processes. Advances in Applied Probability 37, 342–365.

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

Holz, F., Richter, P.M., Egging, R., 2016. The role of natural gas in a low-carbon Europe: Infrastructure and supply security. Energy Journal 37, 33–59.

Hong, H., Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? Journal of Financial Economics 105, 473 – 490.

Hoogerheide, L., Opschoor, A., van Dijk, H., 2012. A Class of adaptive importance sampling weighted EM algorithms for efficient and robust posterior and predictive simulation. Tinbergen Institute Discussion Papers 12-026/4. Tinbergen Institute.

Hou, C., Nguyen, B.H., 2018. Understanding the U.S. natural gas market: A Markov switching VAR approach. Energy Economics 75, 42–53.

Hudson, E.A., Jorgenson, D., 1974. U.S. Energy Policy and Economic Growth, 1975-2000. Bell Journal of Economics 5, 461–514.

Hunt, L., Judge, G., Ninomiya, Y., 2003. Underlying trends and seasonality in UK energy demand: A sectoral analysis. Energy Economics 25, 93–118.

Huurman, C., Ravazzolo, F., Zhou, C., 2012. The power of weather. Computational Statistics & Data Analysis 56, 3793 – 3807.

Jacquier, E., Polson, N.G., Rossi, P., 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics 122, 185–212.

Jadidzadeh, A., Serletis, A., 2017. How does the U.S. natural gas market react to demand and supply shocks in the crude oil market? Energy Economics 63, 66–74.

Jevons, W.S., 1865. The Coal Question; An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal Mines. London: Macmillan and Co.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics And Control 12, 231 – 254.

Johansen, S., Juselius, K., 1990. Maximum likelihood estimation and inference on cointegration - with applications to demand for money. Oxford Bulletin of Economics And Statistics 52, 169–210.

Juvenal, L., Petrella, I., 2015. Speculation in the oil market. Journal of Applied Econometrics 30, 621–649.

Kaldor, N., 1939. Speculation and Economic Stability. The Review of Economic Studies 7, 1–27.

Kallsen, J., Tankov, P., 2006. Characterization of dependence of multidimensional lévy processes using lévy copulas. Journal of Multivariate Analysis 97, 1551 – 1572.

Karakatsani, N.V., Bunn, D.W., 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., D'haeseleer, W., 2011. Gas market distorting effects of imbalanced gas balancing rules: Inefficient regulation of pipeline flexibility. Energy Policy 39, 865 – 876.

Kilian, L., 2008a. The economic effects of energy price shocks. Journal of Economic Literature 46, 871–909.

Kilian, L., 2008b. Exogenous oil supply shocks: How big are they and how much do they matter for the u.s. economy? The Review of Economics and Statistics 90, 216–240.

Kilian, L., 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. The American Economic Review 99, 1053–1069. URL: http://www.jstor.org/stable/25592494.

Kilian, L., 2016. The impact of the shale oil revolution on U.S. oil and gasoline prices. Review of Environmental Economics and Policy 10, 185–205.

Kilian, L., 2019. Measuring global real economic activity: Do recent critiques hold up to scrutiny? Economics Letters 178, 106 – 110.

Kilian, L., Lee, T.K., 2014. Quantifying the speculative component in the real price of oil: The role of global oil inventories. Journal of International Money and Finance 42, 71–87.

Kilian, L., Murphy, D.P., 2012. Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market var models. Journal of the European Economic Association 10, 1166–1188.

Kilian, L., Murphy, D.P., 2014a. The role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics 29, 454–478.

Kilian, L., Murphy, D.P., 2014b. The role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics 29, 454–478.

Kilian, L., Vega, C., 2011. Do Energy Prices Respond to U.S. Macroeconomic News? A Test of the Hypothesis of Predetermined Energy Prices. The Review of Economics and Statistics 93, 660–671.

Kilian, L., Zhou, X., 2020. The Econometrics of Oil Market VAR Models. CESifo Working Paper Series 8153. CESifo.

Knetsch, T.A., 2007. Forecasting the price of crude oil via convenience yield predictions. Journal of Forecasting 26, 527–549.

Knittel, C.R., Roberts, M.R., 2005. An empirical examination of restructured electricity prices. Energy Economics 27, 791 – 817.

Korobilis, D., 2013. VAR Forecasting Using Bayesian Variable Selection. Journal of Applied Econometrics 28, 204–230.

Krivine, H., Ameisen, J., 2018. Comprendre sans prévoir, prévoir sans comprendre. Cassini.

Kumar, U., Jain, V., 2010. Time series models to forecast energy consumption in India. Energy 35, 1709–1716.

Kurmann, A., Sims, E., 2020. Revisions in utilization-adjusted tfp and robust identification of news shocks. Forthcoming: The Review of Economics and Statistics 0, 1–45.

Kuttner, K.N., 2001. Monetary policy surprises and interest rates: Evidence from the fed funds futures market. Journal of Monetary Economics 47, 523 – 544.

Känzig, D.R., 2019. The macroeconomic effects of oil supply news: Evidence from opec announcements.

Labys, W.C., 1982. A Critical Review of International Energy Modelling Methodologies. MIT Energy Labatory Working Paper.

Lanne, M., Luoto, J., 2016. Noncausal bayesian vector autoregression. Journal of Applied Econometrics 31, 1392–1406.

Lanne, M., Luoto, J., 2017. A new time-varying parameter autoregressive model for U.S. inflation expectations. Journal of Money, Credit and Banking 49, 969–995.

Lanne, M., Luoto, J., Saikkonen, P., 2012. Optimal forecasting of noncausal autoregressive time series. International Journal of Forecasting 28, 623–631.

Lanne, M., Saikkonen, P., 2008. Modeling expectations with noncausal autoregressions. MPRA Paper 8411. University Library of Munich, Germany.

Lanne, M., Saikkonen, P., 2013a. Noncausal vector autoregression. Econometric Theory 29, 447–481.

Lanne, M., Saikkonen, P., 2013b. Noncausal vector autoregression. Econometric Theory 29, 447–481.

Lee, C.C., Chang, C.P., 2005. Structural breaks, energy consumption, and economic growth revisited: Evidence from Taiwan. Energy Economics 27, 857–872.

Leeper, E.M., Walker, T.B., Yang, S.C.S., 2013. Fiscal foresight and information flows. Econometrica 81, 1115–1145.

Lewbel, A., Ng, S., 2005. Demand systems with nonstationary prices. Review of Economics and Statistics 87, 479–494.

Lippi, F., Nobili, A., 2012. Oil and the macroeconomy: A quantitative structural analysis. Journal of the European Economic Association 10, 1059–1083.

Lippi, M., Reichlin, L., 1994. Var analysis, nonfundamental representations, blaschke matrices. Journal of Econometrics 63, 307–325.

Litterman, R.B., 1986. Forecasting with bayesian vector autoregressions: Five years of experience. Journal of Business & Economic Statistics 4, 25–38.

Lof, M., 2013. Noncausality and asset pricing. Studies in Nonlinear Dynamics & Econometrics 17, 211–220.

Lucia, J.J., Schwartz, E.S., 2002. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. Review of Derivatives Research 5, 5–50.

Mackay, R., Probert, S., 1995. Crude oil and natural gas supplies and demands up to the year 2010 for France". Applied Energy 50, 185 – 208.

Malanima, P., 2012. The path towards the modern economy the role of energy. From Malthus' stagnation to sustained growth: Social, demographic and economic factors. Macmillan UK. pp. 71–99.

Malanima, P., 2014. Energy in History. pp. 1–29.

Malm, A., 2013. The origins of fossil capital: From water to steam in the british cotton industry. Historical Materialism 21, 15–68.

Manne, A.S., Roland, K., Stephan, G., 1986. Security of supply in the Western European market for natural gas. Energy Policy 14, 52–64.

Markandya, A., Pemberton, M., 2010. Energy security, energy modelling and uncertainty. Energy Policy 38, 1609–1613.

McAvinchey, I.D., Yannopoulos, A., 2003. Stationarity, structural change and specification in a demand system: The case of energy. Energy Economics 25, 65–92.

Medlock, K., 2009. Energy demand theory, in: International Handbook on the Economics of Energy. Edward Elgar Publishing. chapter 5.

de Menezes, L.M., Houllier, M.A., 2016. Reassessing the integration of European electricity markets: A fractional cointegration analysis. Energy Economics 53, 132–150.

Miriello, C., Polo, M., 2015. The development of gas hubs in Europe. Energy Policy 84, 177–190.

Moussa, Z., Thomas, A., 2020. A structural non-causal var model of the global oil market: the role of oil supply news shocks. Mimeo .

Moussa, Z., Thomas, A., Sévi, B., 2020. Considering real-time demand to forecast the u.s. natural gas price in real-time: The role of temperature data. Mimeo .

Müller, J., Hirsch, G., 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. p. 109–128.

Nelimarkka, J., 2017a. Evidence on News Shocks under Information Deficiency. MPRA Paper 80850. University Library of Munich, Germany. URL: https://ideas.repec.org/p/pra/mprapa/80850.html.

Nelimarkka, J., 2017b. The effects of government spending under anticipation: the noncausal VAR approach. MPRA Paper 81303. University Library of Munich, Germany. URL: https://ideas.repec.org/p/pra/mprapa/81303.html.

Nguyen, B.H., Okimoto, T., 2019. Asymmetric reactions of the U.S. natural gas market and economic activity. Energy Economics 80, 86–99.

Nguyen, H.T., Nabney, I.T., 2010. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. Energy 35, 3674–3685.

Pal, D., Mitra, S.K., 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.

Paltsev, S., Jacoby, H.D., Reilly, J., Ejaz, Q.J., Morris, J., O'Sullivan, F., Rausch, S., Winchester, N., Kragha, O., 2011. The future of U.S. natural gas production, use, and trade. Energy Policy 39, 5309–5321.

Panagiotidis, T., Rutledge, E., 2007. Oil and gas markets in the U.K.: Evidence from a cointegrating approach. Energy Economics 29, 329–347.

Parikh, J., Purohit, P., Maitra, P., 2007. Demand projections of petroleum products and natural gas in India. Energy 32, 1825–1837.

Paul, P., 2020. The time-varying effect of monetary policy on asset prices. The Review of Economics and Statistics 102, 690–704.

Pesaran, M.H., Shin, Y., 1999. An autoregressive distributed-lag modelling approach to cointegration analysis. Cambridge University Press. Econometric Society Monographs, p. 371–413.

Pesaran, M.H., Shin, Y., Smith, R.J., 2001. Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics 16, 289–326.

Pindyck, R.S., 1994. Inventories and the Short-Run Dynamics of Commodity Prices. RAND Journal of Economics 25.

Pindyck, R.S., 2004. Volatility and commodity price dynamics. Journal of Futures Markets 24, 1029–1047.

Primiceri, G.E., 2005. Time varying structural vector autoregressions and monetary policy. The Review of Economic Studies 72, 821–852.

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

Ramanathan, R., Engle, R., Granger, C.W., Vahid-Araghi, F., Brace, C., 1997. Short-run forecasts 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.

Rosendahl, K.E., Sagen, E.L., 2009. The global natural gas market: Will transport cost reductions lead to lower prices? The Energy Journal 30, 17–39.

Routledge, B.R., Seppi, D.J., Spatt, C.S., 2000. Equilibrium forward curves for commodities. The Journal of Finance 55, 1297–1338.

Ryan, D.L., Plourde, A., 2009. Empirical Modelling of Energy Demand, in: Evans, J., Hunt, L.C. (Eds.), International Handbook on the Economics of Energy. Edward Elgar Publishing. chapter 6.

Schwartz, E.S., 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. Journal of Finance 52, 923–973.

Serletis, A., Herbert, J., 1999. The message in North American energy prices. Energy Economics 21, 471 – 483.

Sharma, D.P., Chandramohanan Nair, P.S., Balasubramanian, R., 2002. Demand for commercial energy in the state of Kerala, India: An econometric analysis with medium-range projections. Energy Policy 30, 781–791.

Shin, Y., Yu, B., Greenwood-Nimmo, M., 2014. Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework. Springer. pp. 281–314.

Smil, V., 1994. Energy In World History. Essays in world history, Avalon Publishing.

Smil, V., 2006. Energy: A Beginner's Guide. Oneworld Publications.

Smil, V., 2015. Natural Gas: Fuel for the 21st Century. Wiley.

Smirnov, N., 1948. Table for estimating the goodness of fit of empirical distributions. Annals of Mathematical Statistics 19, 279–281.

Smith, C., Hall, S., Mabey, N., 1995. Econometric modelling of international carbon tax regimes. Energy Economics 17, 133–146.

Suganthi, L., Samuel, A.A., 2012. Energy models for demand forecasting—a review. Renewable and Sustainable Energy Reviews 16, 1223–1240.

Sun, X., Huang, D., Wu, G., 2012. The current state of offshore wind energy technology development. Energy 41, 298 – 312.

Susan Yang, S.C., 2005. Quantifying tax effects under policy foresight. Journal of Monetary Economics 52, 1557–1568.

Tang, K., Xiong, W., 2012. Index investment and the financialization of commodities. Financial Analysts Journal 68, 54–74.

Thoenes, S., 2014. Understanding the determinants of electricity prices and the impact of the German nuclear moratorium in 2011. Energy Journal 35, 61–78.

Thomas, A., 2020. The role of expectations in predicting the real price of oil: A non causal analysis. Mimeo .

Thomas, A., Massol, O., Sévi, B., 2019. How are day-ahead prices informative for predicting the next day's consumption of natural gas? evidence from france. Mimeo .

Timmer, R.P., Lamb, P.J., 2007. Relations between temperature and residential natural gas consumption in the central and eastern United States. Journal of Applied Meteorology and Climatology 46, 1993–2013.

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

Uhlig, H., 2004. Do technology shocks lead to a fall in total hours worked? Journal of the European Economic Association 2, 361–371.

Vany, A.D., Walls, W.D., 1993. Pipeline access and market integration in the natural gas industry: Evidence from cointegration tests. Energy Journal 14, 1–20.

Von Hirschhausen, C., Neumann, A., 2008. Long-term contracts and asset specificity revisited: An empirical analysis of producer–importer relations in the natural gas industry. Review of Industrial Organization 32, 131–143.

Wang, Y., Liu, L., Wu, C., 2020. Forecasting commodity prices out-of-sample: Can technical indicators help? International Journal of Forecasting 36, 666 – 683.

WEO, 2018. World Energy Outlook 2018. Technical Report. International Energy Agency (IEA).

Wonnacott, R., Wonnacott, T., 1979. Econometrics. Wiley series in probability and mathematical statistics: Applied probability and statistics, Wiley.

Working, H., 1949. The theory of price of storage. American Economic Review 39, 1254–1262.

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

Zenghelis, D., Hepburn, C., Aghion, p., Teytelboym, A., 2014. Path-dependence, innovation and the economics of climate change .

Zhou, X., 2020. Refining the workhorse oil market model. Journal of Applied Econometrics 35, 130–140.

DOCTORAT BRETAGNE LOIRE ECONOMIE ET GESTION

UNIVERSITÉ DE NANTES

Titre : L'économétrie de la demande énergétique : identification et prévision

Mots clés : Demande Energétique ; Prévision ; Identification ; VAR Bayésien ; Non causalité

Résumé : La prévention du changement climatique est l'une des priorités de la politique énergétique mondiale qui vise à réduire massivement les émissions de gaz à effet de serre. Face à ces défis, il est frappant de constater que notre connaissance de la modélisation de la demande énergétique demeure imparfaite car elle repose en grande partie sur des travaux empiriques anciens et des méthodologies aujourd'hui dépassées. L'objectif scientifique de cette thèse est double : analyser quantitativement les déterminants économiques de la demande énergétique et développer de nouveaux modèles de prévision. Cette thèse est structurée en quatre chapitres. Le premier chapitre montre que la consommation de gaz naturel en France peut être prédite à l'aide d'un modèle simple utilisant seulement les informations disponibles pour les acteurs du marché. Ce chapitre prouve l'existence d'une relation à long terme entre la demande de gaz naturel et les prix des autres énergies et il fournit des estimations de leurs impacts marginaux sur les niveaux de demande observés. Le deuxième chapitre étudie empiriquement le rôle de la température dans la prévision des prix du gaz aux États-Unis. Il développe une méthodologie de construction d'un nouvel indice mensuel basé sur la température. Cet indice capture les variations de la demande résiduelle de gaz naturel en temps réel. Il est utilisé comme variable exogène supplémentaire dans des modèles structurels VAR afin d'améliorer les prévisions ; et nous montrons que ces modèles prédictifs dérivés de modèles structurels sont améliorés en s'appuyant sur des données en temps réelles (non sujettes à révision). Le troisième chapitre propose d'utiliser dans le cas du pétrole, un modèle structurel capturant les anticipations à l'aide de VAR non causaux et d'identifier correctement les réactions des variables clés du pétrole à un choc d'actualité. Le quatrième chapitre réexamine le pouvoir prédictif de la structure par terme des prix, dite « convenience yield », du pétrole et du gaz en intégrant les anticipations dans une spécification empirique, par le biais d'un VAR non causal basé sur la théorie du stockage qui fournit des prévisions de prix très compétitives dans un cadre bivarié simple.

Title : The Econometrics of Energy Demand : Identification and Forecast

Keywords : Energy Demand ; Forecast ; Identification ; Bayesian VAR ; Non causality

Abstract : The prevention of climate change is one of the priorities of the world energy policy that aims to massively reduce greenhouse gas emissions. Faced with these challenges, it is striking to note that our knowledge of energy demand modeling remains limited because it is largely based on old empirical work and methodologies that are now dated. Therefore, the objective of our work is twofold. First, we analyze quantitatively the economic determinants of energy demand. Second, we develop new forecasting models. This thesis is structured in four chapters. The first chapter shows that natural gas consumption in France can be predicted using a simple model which only includes public information that is available to market's participants. This chapter proves the existence of a long-term relationship between demand and prices of other energies and provides estimates of their marginal impacts on observed demand levels. The second chapter empirically investigates the role of temperature in forecasting gas prices in the US. It develops a methodology to build a new monthly index based on temperature. This index captures variations in residual demand for natural gas in real time. It is used as an additional exogenous variable in structural models (VAR) to improve forecasts and we show that, in our case, predictive models derived from a structural model are enhanced relying on true real-time (not subject to revisions) data. The third chapter proposes to use, in the case of oil market, a structural model capturing expectations in a non-causal VAR framework, and to properly identify the reactions of oil key variables to supply news shock. The fourth chapter revisits the predictive power of oil and gas convenience yield by incorporating expectations into an empirical specification through non-causal VAR based on the theory of storage which delivers very competitive price predictions in a simple bivariate setting.