Market power and spatial arbitrage between interconnected gas hubs☆

Olivier MASSOL [a,b,c*] **Albert BANAL-ESTAÑOL** [c,d]

January 8, 2018

Abstract

This paper examines the performance of the spatial arbitrages carried out between two regional markets for wholesale natural gas linked by a pipeline system. We develop a new empirical methodology to (i) detect if these markets are integrated, i.e., if all the spatial arbitrage opportunities between the two markets are being exploited, and (ii) decompose the observed spatial price differences into factors such as transportation costs, transportation bottlenecks, and the oligopolistic behavior of the arbitrageurs. Our framework incorporates a new test for the presence of market power and it is thus able to distinguish between physical and strategic behavior constraints on marginal cost pricing. We use the case of the "Interconnector" pipeline linking Belgium and the UK as an application. Our empirical findings show that all the arbitrage opportunities between the two zones are being exploited but confirm the presence of market power.

Keywords: Law of one price, market integration, spatial equilibrium, interconnectors, natural gas.

JEL Classification: D43; F14; F15; L95; Q37.

☆ This paper has greatly benefited from the judicious comments of two anonymous referees. We are also greatly indebted to Michel Le Breton, Derek Bunn and Steven Gabriel for insightful suggestions. We have also benefited from helpful discussions with Frédéric Lantz, Yves Smeers, conference participants at ISEFI (Paris, 2016) and the 9th Conference on Energy Markets (Toulouse, 2014) and seminar participants at the universities of Cambridge, Paris-Dauphine and Grenoble. Of course, any remaining errors are ours. This research has been supported by the French Energy Council (Conseil Français de l'Énergie) under grant CFE-73 and the Chair "The Economics of Natural Gas". The views expressed herein are strictly those of the authors and should not be construed to reflect the opinions of IFP Énergies Nouvelles or the French Energy Council.

a IFP Énergies Nouvelles, 1–4 av. de Bois Préau, F-92852 Rueil-Malmaison, France.

b Center for Economics and Management, IFP School, 228–232 av. Napoléon Bonaparte, F-92852 Rueil-Malmaison, France.

c Department of Economics, City, University of London, Northampton Square, London EC1V 0HB, UK.

d Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain.

* Corresponding author. Tel.: +33 1 47 52 68 26; fax: +33 1 47 52 70 66.
E-mail address: olivier.massol@ifpen.fr

1. Introduction

Over the last two decades, a series of structural and regulatory reforms have been carried out to ensure the competitiveness of the European natural gas industry. A major development in this restructuring was the emergence of a collection of wholesale markets for natural gas, the “gas hubs”, interconnected throughout the pipeline network. Though these hubs were initially developed to cope with local network balancing needs, they turned out to become a source of gas procurement as the previously monopolized industry structure gradually became more fragmented (Miriello and Polo, 2015). Crucially, EU-led reforms also allowed gas arbitrageurs to purchase intermarket transportation rights and compete to exploit spatial price differences between these interconnected markets.

These spatial arbitrages are central to ensuring an efficient supply of natural gas, especially in the EU. Indeed, the policy debates related to the organization of the EU’s internal market for natural gas repeatedly underline the importance of spatial arbitrages as a means to prevent balkanization (Vazquez et al., 2012). The security of the supply of natural gas is a recurrent source of concern in Europe which is predominantly served by a small oligopoly of foreign producers (Abada and Massol, 2011). Geographically, each of these producers can be viewed as a dominant player in the adjacent national markets it serves (e.g., Russia in Eastern European countries, Algeria in Southern Europe). As the producers’ operations are located outside the EU jurisdiction, they can exert market power in supplying these markets, a situation that may be hard to moderate using the usual EU competition arsenal. To overcome that problem, the EU policy strongly promotes the “creation of a single gas market” – i.e., the spatial integration of the national wholesale markets – aimed at diluting the concentration observed in some countries within a wider economic market.

Defining and measuring spatial integration, though, is not straightforward. In Stigler and Sherwin (1985), two geographical markets for a tradable good are set to be integrated if the spatial price difference between these two markets equals the unit transportation cost. However, empirically, assessing the spatial integration of wholesale gas markets remains a challenging task because price spreads could also reflect other factors, including transportation bottlenecks and, more importantly, oligopolistic pricing by the arbitrageurs. To overcome this problem, we define integration using the equilibrium notion that all spatial arbitrage opportunities between the two markets are being exploited, i.e: that price spreads are consistent with the traders’ profit maximization behavior. This notion is derived from the theoretical literature on spatial price determination that was pioneered by Enke (1951), Samuelson (1952), Takayama and Judge (1971) and Harker (1986).

This paper develops a new empirical methodology to assess the arbitrages between two regional markets for wholesale natural gas linked by a capacity-constrained infrastructure. This methodology is designed to (i) detect if these markets are integrated, i.e., if all the spatial arbitrage opportunities are being exploited, and (ii) decompose the observed spatial price differences into factors such as transportation costs, transportation bottlenecks, and the oligopolistic behavior of the arbitrageurs. Our framework incorporates a test for the presence of market power and is thus able to distinguish between physical and strategic constraints to marginal cost pricing. As an application, we use the spatial arbitrages in the "Interconnector" pipeline which connects Europe's two oldest spot markets for natural gas: the UK's National Balancing Point and the Zeebrugge market in Belgium.

A large amount of empirical research has examined the degree of spatial integration between markets for wholesale natural gas with the help of time-series techniques. These studies typically rely on local price data and assess the co-movements of prices at each market location. In these analyses, it is typically argued that high degrees of correlation (Doane and Spulber, 1994) and/or co-integration between the price series (e.g., De Vany and Walls, 1993; Serletis, 1997; Asche et al., 2002; Siliverstovs et al., 2005; Asche et al., 2013) are evidence that the law of one price is being enforced through spatial arbitrages.[1] These price-based empirical models provide useful insights into how local price shocks are transmitted to adjacent markets. However, the methodology used in these studies is of little help in assessing the competitive nature of the observed spatial arbitrages, as they fail to detect the presence of imperfect competition. Moreover, as suggested by Barrett (1996, 2001), Baulch (1997) and McNew and Fackler (1997), these empirical models are unable to account for the pivotal role played by both intermarket transfer costs and trade flow considerations.[2]

This paper uses, and brings to the field of energy economics for the first time, an alternative approach based on the extended parity bounds model (PBM) developed by Barrett and Li (2002). In a PBM, arbitrageurs are assumed to be profit-maximizing agents. Using that assumption, intermarket price spreads are examined using a "switching regime" specification, which estimates the probability of observing each of a series of trade regimes. Sexton et al. (1991), for example, use only price data and consider three distinct trade regimes depending on whether the spatial price difference is greater, equal or lower than the unit intermarket transportation cost. This modeling approach is now widely used in agricultural economics to assess food market integration (Baulch, 1997; Fackler and Goodwin, 2001; Negassa and Myers, 2007; Cirera and Arndt, 2008; Moser et al., 2009; Zant, 2013). In particular, Barrett and Li (2002), our point of departure, make use of trade flow data to further distinguish whether trade occurs or not in each of the three regimes. Their direction-specific approach allows them to detect any violation of the theoretical equilibrium conditions that all arbitrage opportunities between the two markets are being exploited: namely, if trade is observed and the spatial

[1] Other time-series analyses include (i) the autoregressive model of pairwise price differentials in Cuddington and Wang (2006) to estimate the speed of adjustment toward equilibrium, (ii) the vector error-correction models in Park et al. (2008), Brown and Yücel (2008) or Olsen et al. (2015), and (iii) the examinations a time-varying degree of price convergence among spot markets with the help of the Kalman Filter approach (King and Cuc, 1996; Neumann et al., 2006; Renou-Maissant, 2012).

[2] These criticisms emphasize a lack of acquaintance with existing economic models of spatial price determination. Two lines of arguments motivate that shortcoming. First, intermarket transfer costs are typically omitted in these early empirical studies whereas, in theory, price equalizing arbitrage activities are triggered only when localized shocks result in spatial price differences which exceed these intermarket transfer costs (Barrett, 1996, 2001; Baulch, 1997; McNew and Fackler, 1997). Second, trade flows information play no role in these early empirical studies whereas theory suggests that either discontinuities in the trade flows or variations in the directions of these flows can have an impact on the degree of co-movements among prices at each market location (Barrett and Li, 2002).

price difference is lower than the transportation cost or if no trade is observed while having a spatial price spread greater than the transportation cost.[3]

We introduce two modifications to existing PBMs to be able to apply them to natural gas markets. Existing PBMs assume the presence of perfect competition in the spatial arbitrages, which may be a realistic assumption in agricultural markets, but not in natural gas markets. So, we introduce, for the first time in a PBM, the possibility that arbitrageurs have market power. Second, again because of the nature of agricultural markets, the role of transportation bottlenecks has so far been neglected. But binding (pipeline) capacity constraints are highly likely to occur in the gas industry. So, we propose to analyze, for the first time, the role of capacity constraints in spatial price spreads.

As an application, we examine the spatial arbitrages performed between the two oldest European markets for wholesale natural gas in Belgium and the UK. This allows us to present a series of original empirical findings that: (i) show that all the arbitrage opportunities between the two zones are being exploited, but (ii) confirm the presence of market power in the spatial arbitrages. As the detailed institutional arrangements created for these two markets have largely shaped the designs of the other Continental markets, we believe that these findings provide a valuable contribution to the policy debate related to the restructuring of the European market for natural gas.

We believe that this framework can provide useful guidance to a large audience interested in the functioning of the restructured natural gas industries (e.g., competition authorities, regulators, market analysts). It could also inform the modeling choices retained in the spatial equilibrium models recently developed for that industry (cf., Huntington, 2009). In these numerical models, researchers typically either posit the existence of competitive spatial arbitrages (e.g., Golombek et al., 1995) or imperfectly competitive ones (e.g., Abada et al., 2013).

Despite the importance of market power concerns in the energy policy debates, the market power potentially exerted by natural gas arbitragers has hitherto been little studied. A notable exception is the theoretical analysis in Ritz (2014) who highlights the potential role of the LNG exporters' market power in the observed price differentials between Asia and Northwest Europe. In an empirical analysis, Rupérez Micola and Bunn (2007) apply standard regression techniques to examine the relationship between the pipeline capacity utilization (i.e., the ratio of utilized to maximum capacity) and the absolute price difference between Belgium and the UK. Their results document the presence of market splitting at moderate levels of capacity utilization which, according to the authors, suggests the presence of market power inefficiencies. However, neither the direction of the trade flows nor the intermarket transfer costs play any role in their analysis. By taking these features into account, our

[3] The PBM framework was first proposed by Sexton et al. (1991), drawing on Spiller and Huang (1986). To the best of our knowledge, the energy economics literature only provides a handful of studies based on the PBM framework (e.g., Kleit, 1998, 2001; Bailey, 1998). All of them only use price data, are based on the original work of Spiller and Huang (1986) and ignore the developments proposed in Barret and Li (2002). Hence, these earlier models are unable to test whether all arbitrage opportunities between the two markets are being exploited.

paper confirms the presence of market power, even if all the arbitrage opportunities are being exploited, and connects the empirical results to the theoretical literature on spatial price determination.

The remaining sections of this paper are organized as follows. Section 2 details the theoretical conditions for spatial equilibrium between two markets linked by a capacity-constrained transportation infrastructure. Section 3 presents an adapted empirical methodology to investigate whether these conditions hold or not. Then, Section 4 details an application of this methodology to the case of the Interconnector UK, a natural gas pipeline connecting the UK to Continental Europe. Finally, the last section offers a summary and some concluding remarks.

2. Theoretical background

This section presents the theoretical conditions for short-run spatial equilibrium between two markets connected by a capacity-constrained transportation infrastructure. We make two alternative assumptions regarding the *gas traders*' behavior: perfect competition and oligopolistic behavior à la Cournot. For simplicity, in both cases, the *local* supply in both markets is assumed to be competitive.

We consider two markets i and j located in different regions that trade a homogeneous commodity. At time t, the two markets are connected by a single transportation infrastructure that has direction-specific, non-negative finite capacities that can change over time, K_{jit} from market from j to i and K_{ijt} in the opposite direction. During any trading period t, the infrastructure solely allows to move the commodity in a given direction, and therefore the directional capacities *de facto* verify $K_{jit} \times K_{ijt} = 0$. The direction of the infrastructure is known at the beginning of each trading period. It can vary from one such period to the next but not within a given trading period. This framework is consistent with the physical constraints observed in the natural gas industry.[4]

Let us, from now on, concentrate on the direction-specific arbitrages that can be performed from market j to market i at time t. For each region i at time t, we assume that there is a linear inverse demand function: $p_{it}^{D}(q) = a_{it} - b_i q$ with $a_{it} > 0$ and $b_i > 0$.[5] We assume that the local industries' aggregate supply functions are linear and upward sloping.[6] In each region i at time t, we let

[4] A bidirectional pipeline can only be operated in a given predetermined direction during the normal trading hours of a given day. Because of substantial system inertia, the flow direction cannot be changed within the trading hours. Therefore, the direction of the infrastructure is typically communicated to traders at the beginning of the trading hours who, in turn, know that their arbitrage decisions cannot modify the direction of the infrastructure within the day.

[5] These slope coefficients are not subscripted with the time index and are thus assumed to be constant. In contrast, the intercepts of these inverse demand functions are assumed to be time-varying parameters (because of the seasonal variations observed in natural gas demand). These assumptions are frequently used in the context of restructured electricity markets (e.g., Day and Bunn, 2001).

[6] The use of a linear functional form for the demand and supply curves simplifies computations and is commonly adopted in the energy economics literature (e.g., Green, 1999). That said, our main arguments should hold even if demand and supply functions were not linear because in the presence of market power, each arbitrageur recognizes that the quantity it supplies will affect the prices at both the originating and destination markets. As a result, the price paid and the price received will depend on the quantity supplied, and the resulting gains (spread multiplied by quantity) will be non-linear in the quantity supplied. Instead, in the competitive case, the spread gains would be linear in the quantity transported.

$p_{it}^{S}\left(S_{it}\right)=c_{it}+d_i\,S_{it}$, with $c_{it}>0$ and $d_i>0$ and where S_{it} is the local supply, denote the inverse supply function. To avoid corner solutions, we also assume that: (i) in the destination market i, the price demanded when the import infrastructure is operated at full capacity is larger than the marginal cost of the least-costly local supplier (i.e., $a_{it}-b_i\,K_{jit}>c_{it}$), and (ii) in the origin market j, the marginal cost to solely supply a flow equal to the infrastructure capacity is lower than the local consumers' highest willingness to pay for that commodity (i.e., $a_{jt}>c_{jt}+d_jK_{jit}$).

The trading firms' unique activity is to perform spatial arbitrages.[7] We assume that there are no transport lags so that spatial arbitrage can take place within each observation period. The non-negative aggregate trade flow from j to i measured at time t is denoted Q_{jit}.

When performing a spatial arbitrage from market j to market i at time t, a trader incurs the unit transfer costs T_{jit}. In addition, that trader must also purchase an appropriate amount of transportation rights. A transportation right provides its owner with the right to transfer up to one unit of good from market j to market i at a given time period t. We let ξ_{jit} denote the market clearing price of a transportation right from market j to market i at time t. The total number of rights offered at that time is K_{jit}.

a – Case A: Perfectly competitive spatial arbitrages

In this case, we assume that traders adopt a price-taking behavior at each location and in the market for transportation rights. Following the logic of the Enke-Samuelson-Takayama-Judge spatial equilibrium model, the traders' aggregate behavior at time t can be described using two complementarity conditions that together characterize the equilibrium conditions for competitive spatial arbitrages:

Proposition 1*: For the equilibrium conditions for perfectly spatial arbitrages to hold at time t, we need that the following aggregate complementarity conditions are verified:*

$$0\le Q_{jit}, \qquad P_{it}-P_{jt}-T_{jit}-\xi_{jit}\le 0 \qquad \textit{and} \qquad \left(P_{it}-P_{jt}-T_{jit}-\xi_{jit}\right)Q_{jit}=0, \tag{1}$$

$$0\le \xi_{jit}, \qquad Q_{jit}\le K_{jit} \qquad \textit{and} \qquad \left(Q_{jit}-K_{jit}\right)\xi_{jit}=0, \tag{2}$$

where P_{it} and P_{jt} are the market clearing prices in each location, and ξ_{jit} is the price of a transportation right that is the price of capacity (in excess of T_{jit}) that ensures that demand for transportation services does not exceed supply K_{jit}.

[7] *Throughout the paper, we follow the convention retained in most of the literature in industrial and energy economics and assume that arbitragers are risk-neutral (e.g., Hubbard and Weiner, 1986; Borenstein et al., 2008).*

The proof of that proposition is detailed in Appendix A. The complementarity condition (1) clarifies the value of the traders' marginal profit to spatial arbitrages. This marginal profit is defined as the difference between the market-clearing price at location i and the marginal cost which is the sum of three elements: the price at location j, the marginal transfer cost T_{jit} and ξ_{jit} the price of a transportation right. The complementarity condition (2) ensures that the price ξ_{jit} is equal to zero whenever the aggregate demand for transportation right Q_{jit} is lower than the supply (i.e., whenever the transportation capacity constraint $Q_{jit} \leq K_{jit}$ is slack), and that ξ_{jit} is positive when the constraint $Q_{jit} \leq K_{jit}$ is binding. In case of a zero price for transportation right (i.e., $\xi_{jit} = 0$), the complementarity condition (1) ensures: (i) that there is no trade from market j to market i (i.e., $Q_{jit} = 0$) when the marginal profit to spatial arbitrage is negative, and (ii) that the spatial price spread is equal to the marginal transfer cost T_{jit} when trade occurs and it is not constrained by the infrastructure's capacity (i.e., $0 < Q_{jit} < K_{jit}$). In case of a binding capacity constraint (i.e., $Q_{jit} = K_{jit}$), the complementary condition (1) ensures that the spatial price difference is larger than the marginal transfer cost T_{jit}. In this case, there exists a scarcity rent $\xi_{jit} K_{jit} = \left(P_{it} - P_{jt} - T_{jit}\right) K_{jit}$.

b – Case B: Oligopolistic spatial arbitrages

We now assume that there are a total of G gas traders that behave à la Cournot in the local markets. As in Harker (1986), each trader thus knows how the prices in each region react to the quantities supplied and demanded thorough the intermarket infrastructure but rather takes the price of transportation right as given. The following proposition indicates that the traders' aggregate behavior at time t can also be described using two complementarity conditions.

Proposition 2: *For the equilibrium conditions for oligopolistic spatial arbitrages to hold at time t, we need that the following aggregate complementarity conditions are verified:*

$$\begin{cases} 0 \leq Q_{jit}, \qquad P_{it} - P_{jt} - T_{jit} - \left(\dfrac{d_i b_i}{b_i + d_i} + \dfrac{d_j b_j}{b_j + d_j} \right) \dfrac{Q_{jit}}{G} - \xi_{jit} \leq 0 \qquad \textit{and} \\ \left(P_{it} - P_{jt} - T_{jit} - \left(\dfrac{d_i b_i}{b_i + d_i} + \dfrac{d_j b_j}{b_j + d_j} \right) \dfrac{Q_{jit}}{G} - \xi_{jit} \right) Q_{jit} = 0, \end{cases} \tag{3}$$

$$0 \leq \xi_{jit}, \qquad Q_{jit} \leq K_{jit} \qquad \text{and} \qquad \left(Q_{jit} - K_{jit}\right) \xi_{jit} = 0, \tag{4}$$

where P_{it} and P_{jt} are the local market clearing prices.

The proof of that proposition is detailed in Appendix B. The economic interpretation of these conditions is similar to those detailed for the case of competitive arbitrages except that the traders' aggregate behavior at time t now accounts for the players' ability to exert market power in both markets. At equilibrium, the marginal revenue obtained by a trader in the destination market is $P_{it} - \frac{d_i b_i}{b_i + d_i}.\frac{Q_{jit}}{G}$ where $-\frac{d_i b_i}{b_i + d_i}.\frac{Q_{jit}}{G}$ represents the marginal loss of revenue from getting a lower price for each of the units the player is selling there. The trader's marginal cost includes three distinct components: (i) the marginal purchase cost in the origin market $P_{jt} + \frac{d_j b_j}{b_j + d_j}.\frac{Q_{jit}}{G}$ where $\frac{d_j b_j}{b_j + d_j}.\frac{Q_{jit}}{G}$ represents the marginal cost increase from getting a higher price for each of the units the player is purchasing (i.e., the effect of the players' oligopsonistic behavior in market j); (ii) the marginal transfer cost T_{jit}, and (iii) the price of a transportation right ξ_{jit}.

It is instructive to compare these aggregate conditions with the ones obtained in case of competitive arbitrage. One may remark that, in the event of oligopolistic arbitrages, the spatial price differential in (3) is always larger than the marginal transfer cost T_{jit} when trade is observed. Moreover, from an empirical perspective, it is interesting to note that the spread between the spatial price differential and the marginal transfer cost is proportional to the aggregate trade flow when trade is observed and the congestion constraint is slack. We shall come back to this point in the sequel. For the moment, we simply highlight that this reflects the traders' ability to exert market power by restricting intermarket trade to generate some oligopolistic rents.

3. Methodology

This section presents the methodology used in this manuscript. We first adapt the existing PBM framework to take into account the role of both pipeline capacity constraints and market power. Subsequently, we detail the empirical specification.

3.1 An adapted parity bounds model

We now define seven mutually exclusive trade regimes and relate them to the theoretical conditions for spatial equilibrium detailed in the previous section. In addition to the six trade regimes considered in the PBM proposed in Barrett and Li (2002), we introduce a new one that takes into account the case of pipeline congestion. Moreover, for each of these trade regimes, we distinguish between the cases of perfectly competitive and oligopolistic spatial arbitrages.

From an empirical perspective, it is important to highlight that the price of a transportation right ξ_{jit} is seldom publicly available as its formation chiefly results from over the counter transactions. In the sequel, we thus follow the convention in Barrett and Li (2002) and define the marginal rents to spatial arbitrage in case of competitive arbitrages as the difference between the spatial price spread and the marginal transfer cost T_{jit}. For ease of exposition, we also define the marginal rents to spatial arbitrage in case of oligopolistic arbitrages as the difference between the marginal revenue obtained in market i (i.e., including the term representing the traders' ability to exert oligopolistic market power there) and the sum of the marginal purchase cost in market j (i.e., including the term representing the traders' ability to exert oligopsonistic market power there) and the marginal transfer cost T_{jit}. Hence, both definitions do not include the price of a transportation right and thus differs from the marginal profit to spatial arbitrages defined above.

As shown in Table 1, marginal rents to spatial arbitrage and trade flow considerations can be combined to define a taxonomy of trade regimes governing the arbitrages from market j to market i. Regarding marginal rents to spatial arbitrage, three basic states can be defined depending on their value: zero, strictly positive, or strictly negative. Regarding trade flows, two basic states can be identified depending on whether a positive trade flow is observed or not. Following Barrett and Li (2002), each of these six regimes is labeled I to VI, where odd numbers are used for regimes with strictly positive trade flows and even numbers for those without trade.

Table 1. The trade regimes in each direction

	Trade is observed: $0 < Q_{jit} \le K_{jit}$	No trade is observed: $Q_{jit} = 0$
zero marginal rents to spatial arbitrage	Regime I λ_I	Regime II λ_{II}
positive marginal rents to spatial arbitrage	Regime III$_a$ iff $Q_{jit} < K_{jit}$ λ_{III_a} Regime III$_b$ iff $Q_{jit} = K_{jit}$ λ_{III_b}	Regime IV λ_{IV}
negative marginal rents to spatial arbitrage	Regime V λ_V	Regime VI λ_{VI}

In regimes I and II, the marginal rents to spatial arbitrage are equal to 0. As shown in the previous section, depending on the assumption posited for the behavior of the trading sector, one of the following conditions is binding:

Case A: Competitive arbitrages

$$P_{it} - P_{jt} - T_{jit} = 0 \quad (5)$$

Case B: Oligopolistic arbitrages

$$P_{it} - P_{jt} - T_{jit} - \left(\frac{d_i b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j} \right) \frac{Q_{jit}}{G} = 0 \quad (6)$$

In case of price-taking behavior (Case A), the spatial price differential is equal to the marginal transfer cost. In case of oligopolistic arbitrages (Case B), the possibility to exert market power results in a spatial price differential that exceeds the marginal transfer cost and the difference between the two is proportional to the observed trade flow. In Case A (respectively B), each of the two regimes verifies the complementarity slackness condition (1) (respectively (3)) when there is no congestion cost (i.e., $\xi_{jit}=0$). Therefore, both regimes are consistent with the conditions for a spatial equilibrium.

In regimes III and IV, the marginal rents to spatial arbitrage from j to i are strictly positive:

Case A: Competitive arbitrages	Case B: Oligopolistic arbitrages
$P_{it}-P_{jt}-T_{jit}>0$ (7)	$P_{it}-P_{jt}-T_{jit}-\left(\frac{d_i b_i}{b_i+d_i}+\frac{d_j b_j}{b_j+d_j}\right)\frac{Q_{jit}}{G}>0$ (8)

In both of these regimes, markets are separated and there are unseized opportunities for profitable spatial arbitrage. Still, in case of positive trade (regime III), the observed insufficient arbitrages might result from the capacity-constrained nature of the transportation infrastructure. Indeed, the complementarity conditions detailed in the preceding section indicate that, in case of a binding capacity constraint (i.e., $Q_{jit}=K_{jit}$), observing a strictly positive value for the marginal rents to arbitrage is consistent with the conditions for a short-run spatial equilibrium. In contrast, the joint observation of strictly positive marginal rents to arbitrages and a slackening in the infrastructure's capacity constraint violates the conditions for a spatial equilibrium. Thus, we propose a modification to the original model and further decompose regime III into two mutually exclusive regimes labeled III_a and III_b. In regime III_a, the observed trade flows verify $0<Q_{jit}<K_{jit}$ whereas a binding capacity constraint (i.e., $Q_{jit}=K_{jit}$) is observed in regime III_b. Therefore, the latter regime, but not the former, is consistent with the conditions for a spatial equilibrium.

In regimes V and VI, the marginal rents to arbitrage from j to i are strictly negative:

Case A: Competitive arbitrages	Case B: Oligopolistic arbitrages
$P_{it}-P_{jt}-T_{jit}<0$ (9)	$P_{it}-P_{jt}-T_{jit}-\left(\frac{d_i b_i}{b_i+d_i}+\frac{d_j b_j}{b_j+d_j}\right)\frac{Q_{jit}}{G}<0$ (10)

In both regimes, there are no profitable arbitrage opportunities. In regime VI, trade is not occurring and the observed local prices correspond to autarky prices. This regime is consistent with the conditions for a spatial equilibrium. In contrast, regime V indicates that trade is occurring despite negative marginal profits which is not consistent with equilibrium conditions.

In sum, having introduced a further distinction between regimes III_a and III_b, a total of seven regimes are thus considered in our analysis. The estimated probability to observe regime r is denoted λ_r. Spatial equilibrium conditions hold with probability $\left(\lambda_I+\lambda_{II}+\lambda_{III_b}+\lambda_{VI}\right)$ and the estimated probability to observe disequilibrium is $\left(\lambda_{III_a}+\lambda_{IV}+\lambda_V\right)$.

3.2 Empirical specification

We now detail the empirical specification aimed at estimating the probabilities of being in each regime using a data set of N observations for the local market-clearing prices, the marginal transfer cost, the trade flow, and the available transportation capacity.

In case of oligopolistic arbitrages, the slope coefficients for the local inverse demand and supply functions are unlikely to be readily available to the modeler. So, we introduce γ an unknown parameter to be estimated that will be interpreted as $\frac{1}{G}\left(\frac{d_i b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j}\right)$ the sum of the two local coefficients determined by the slopes of the inverse supply and inverse demand functions. So, we expect the estimated value for γ to be non-negative.

Denoting $R_{jit} \equiv P_{it} - P_{jt} - T_{jit}$ the series that represents the difference between the spatial price spread and the unit transfer cost, the marginal rents to arbitrage in each of the three distinct cases (zero, positive and negative) are modeled using the following switching regression model (Sexton et al., 1991; Baulch, 1997; Barrett and Li, 2002):

Case A: Competitive arbitrages		Case B: Oligopolistic arbitrages	
Regimes I & II:		Regimes I & II:	
$R_{jit} = \varepsilon_{jit}$	(11)	$R_{jit} = Q_{jit}\gamma + \varepsilon_{jit}$	(14)
Regimes $\mathrm{III_a}$, $\mathrm{III_b}$ & IV:		Regimes $\mathrm{III_a}$, $\mathrm{III_b}$ & IV:	
$R_{jit} = \varepsilon_{jit} + \mu_{jit}$	(12)	$R_{jit} = Q_{jit}\gamma + \varepsilon_{jit} + \mu_{jit}$	(15)
Regimes V & VI:		Regimes V & VI:	
$R_{jit} = \varepsilon_{jit} - \upsilon_{jit}$	(13)	$R_{jit} = Q_{jit}\gamma + \varepsilon_{jit} - \upsilon_{jit}$	(16)

where: γ is an unbounded parameter to be estimated; ε_{jit} is a random error that is assumed to be i.i.d. normally distributed with a zero mean and variance σ_ε^2; and μ_{jit} and υ_{jit} are i.i.d. random samples from zero-centered normal distributions truncated above at 0 with respective variance parameters σ_μ^2 and σ_υ^2.

In applications, measurement and sampling error are likely to occur with any data to which this model might be applied. An extended specification may be justified to control for these issues. In the sequel, a term $\alpha_{ji} + X_t^{'}\beta_{ji}$ – where: α_{ji} is the regime-invariant mean parameter, X_t is a vector of exogenous factors including a time trend and a list of seasonal dummy variables and β_{ji} is the associated regime-invariant vector of parameters – is systematically implemented in the equations to capture the time-invariant, the trend and the seasonal components of possible measurement errors.

The specifications used to model the cases of competitive and oligopolistic spatial arbitrages differ only in the markup term $Q_{jit}\gamma$. Thus, a statistical test of the null hypothesis $\gamma = 0$ (e.g., a likelihood ratio test) can be conducted to test the null hypothesis of perfectly competitive spatial arbitrages. For the sake of brevity, only the unrestricted model based on equations (14), (15), and (16) is detailed hereafter.

Denoting λ the vector of the probabilities to observe the seven regimes, $\theta \equiv \left(\alpha_{ji}, \beta_{ji}, \gamma, \sigma_{\varepsilon}, \sigma_{\mu}, \sigma_{\upsilon}\right)$ the parameter vector to be estimated and $\pi_{jit} \equiv R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma$ the random variable that gives the marginal profit from spatial arbitrage at time t, the joint density function for the observation at time t is the mixture distribution:

$$\begin{aligned} f_{jit}\left(\pi_{jit} \middle| (\lambda,\theta)\right) \;\equiv\; & A_{jit}\left[\lambda_I f_{jit}^{I}\left(\pi_{jit}\middle|\theta\right) + \left(\left(1-B_{jit}\right)\lambda_{III_a} + B_{jit}\lambda_{III_b}\right) f_{jit}^{III}\left(\pi_{jit}\middle|\theta\right) + \lambda_V f_{jit}^{V}\left(\pi_{jit}\middle|\theta\right)\right] \\ & + \left(1-A_{jit}\right)\left[\lambda_{II} f_{jit}^{II}\left(\pi_{jit}\middle|\theta\right) + \lambda_{IV} f_{jit}^{IV}\left(\pi_{jit}\middle|\theta\right) + \lambda_{VI} f_{jit}^{VI}\left(\pi_{jit}\middle|\theta\right)\right] \end{aligned} \quad (17)$$

where: A_{jit} is an indicator variable that takes a value of 1 if trade is observed and zero otherwise; B_{jit} is an indicator variable that takes a value of 1 if the transportation infrastructure is congested and zero otherwise; $f_{jit}^{I}\left(\pi_{jit}\middle|\theta\right)$ and $f_{jit}^{II}\left(\pi_{jit}\middle|\theta\right)$ are normal density functions; $f_{jit}^{III}\left(\pi_{jit}\middle|\theta\right)$ and $f_{jit}^{IV}\left(\pi_{jit}\middle|\theta\right)$ (respectively $f_{jit}^{V}\left(\pi_{jit}\middle|\theta\right)$ and $f_{jit}^{VI}\left(\pi_{jit}\middle|\theta\right)$) are the density functions derived in Weinstein (1964) for the sum of a normal random variable and a centered-normal random variable truncated above (respectively below) at 0.

The likelihood function for a sample of observations $\left\{R_{jit}, Q_{jit}, K_{jit}\right\}$ is:

$$L(\lambda,\theta) \equiv \prod_{t=1}^{N} f_{jit}\left(\pi_{jit}\middle|(\lambda,\theta)\right) \quad (18)$$

The model can be estimated by maximizing the logarithm of the likelihood function with respect to regime probabilities and model parameters subject to the constraints that the regime probabilities sum to one and that each of these probabilities lies in the unit interval.

As most existing PBM models, the specification described so far is based on a static formulation whereby shocks are posited to be serially independent and the variance parameters are held constant throughout the entire observation period. As these assumptions can be too restrictive in applications based on daily data, we also use an enriched dynamic specification including a correction for serial correlation and GARCH-type time-varying variance, as detailed in Appendix C.

This specification differs from that of Barrett and Li (2002) in three ways. First, we show how a parity bound model can be used to test the null hypothesis of competitive spatial arbitrages. Second, contrary to Barrett and Li, the two markets under scrutiny are connected by a capacity-constrained transportation infrastructure. So, a seventh regime labeled III_b is introduced to account for the explanatory role played by infrastructure congestion issues in the observation of positive marginal profits to spatial arbitrage. Lastly, the specification is extended to a dynamic formulation.

4. Application

4.1 Background

This application focuses on the so-called Interconnector (hereafter abbreviated to IUK), a bi-directional natural gas pipeline system connecting the UK National Transportation System (using the Bacton Terminal) to Zeebrugge (Belgium). This infrastructure allows spatial arbitrages between Europe's two oldest spot markets for natural gas: (i) the UK's NBP, which allows counterparties to trade a standardized lot of natural gas piped via the UK National Transmission System with a delivery point at the so-called National Balancing Point (NBP); and (ii) the Zeebrugge local market in Belgium, which is labeled ZEE.

We consider the period covering October 1, 2003, to October 5, 2006. This starting date has been chosen to omit the number of partial closures of the IUK that happened during the summer of 2003 (Futyan, 2006). This terminal date corresponds to the opening of the Langeled infrastructure, a pipeline system that together with already existing offshore pipelines, allowed Norwegian gas producers to perform spatial arbitrages between the UK and the Continent, thereby offering an alternative to the IUK. During that period, the IUK pipeline was thus the unique infrastructure linking the UK and Continental natural gas markets. From an industrial organization perspective, both countries experienced stable market structures during this period which is posterior to the deep restructuring process of the UK gas sector (Wright, 2006) and precedes the merger between Gaz de France and Suez that strongly impacted the Belgian market after November 2006 (Argentesi et al., 2017). In addition, that period corresponds to a steady institutional environment with unchanged access rules for both the IUK and the adjacent national pipeline systems. These features make the IUK case an attractive experiment to assess the spatial price arbitrages that can be performed in a deregulated natural gas industry.

4.2 Data

We use daily transaction price data for day-ahead wholesale natural gas traded during working days as published by Platt's, a price-reporting service. For each working day (i.e., Monday to Friday), they reflect the price range of a standardized quantity of natural gas to be delivered at a constant flow rate throughout the next working day after assessment (Platt's, 2012). All prices are denominated in €/MWh. Given the extremely limited liquidity of within-day markets, we follow the usual convention and refer to these day-ahead prices as "spot" since they provide traders with a final opportunity to trade gas out of a forward position before physical delivery.

The unit transfer costs T_{jit} are direction-specific and are derived from the fuel used by the IUK operator to power its compressor equipment since this cost is billed to the traders. According to the pipeline operator, fuel gas consumption amounts to 0.8% of the quantity of gas transported when natural gas is piped from the UK to Belgium, and to 0.26% of the quantity of gas transported in the other direction. This fuel cost is evaluated using the price of natural gas in the exporting market.

Regarding trade flow data, the wish may be to use an aggregate variable gathering all the transportation nominations communicated at the end of any working day for delivery during the next working day. Unfortunately, these data are confidential. So, this study uses a proxy: a historical flow series representing the physical daily flow of natural gas, measured in GWh/day, that transited through the IUK as reported on the pipeline operator's website. Thus, we proceed under the assumption that the physical gas flow measured during a given working day represents an unbiased estimator of the aggregate transportation nominations decided during the previous working day (at the time when trade occurs in the corresponding day-ahead market).[8]

According to the Interconnector operator, the nominal transportation capacity from the UK to Belgium remained unchanged during the entire sample period. In the other direction, the installation of some compressor equipment in Zeebrugge on November 8, 2005, increased the transportation capacity. Unfortunately, information related to the available daily transportation capacities remains unavailable. So, we follow Rupérez Micola and Bunn (2007) and consider the historical maximum values of the trade flows. The historical maxima were: 624.63 GWh/day from the UK to Belgium, and 310.24 GWh/d prior to November 8, 2005, (respectively 511.80 GWh/d after that date) in the other direction. Nevertheless, it should be noted that the daily flow capacity of a point-to-point natural gas pipeline is a time-varying parameter that depends on a series of exogenous factors (e.g., the operating pressures of the adjacent national pipeline systems, the flow temperature, the chemical composition of the natural gas (Yépez, 2008; Massol, 2011). Hence, the historical maximum daily flow cannot necessarily be attained. We proceed by assuming that congestion is likely to be a source of concern when the observed capacity utilization ratio (measured against the historically maximum) exceeds 80%.[9] Hereafter, this threshold is used to distinguish regimes III_a and III_b.

The data set has been modified in two ways. First, both the Belgium and UK markets are closed on national bank holidays. To account for differences in the national calendars, all the observations related to a bank holiday in either Belgium or the UK have been disregarded in the subsequent analyses. Second, we excluded observations made on dates during which the Interconnector service was unavailable due to planned maintenance (these dates are documented on the IUK's website). As a

[8] On a given working day, pipeline users are offered the possibility to revise the transportation service requested at the end of the previous working day. This is the so-called within-day re-nominations. Yet, for the pipeline operator, these within-day re-nominations generate a significant extra operation costs. As a result, the detailed pricing rules adopted by the pipeline operator have been explicitly designed to render these within-day re-nominations extremely costly. So, users have a strong incentive to contract their real transportation needs for day d+1 at the end of day d (i.e., before the close of the day ahead market). Therefore, we proceed assuming that these within-day re-nominations can be neglected.

[9] Regarding the threshold level, we also tested two other threshold levels: 85% and 90% but the estimation results were very similar to the ones obtained using the 80% level.

result, we assembled time series data containing 723 daily observations on prices, compressor fuel costs, and trade flows in each direction.

Lastly, the constant mean parameter α_{ji} is supplemented by a vector of observable exogenous variables in order to control for the possible impact of a time-varying measurement bias in the price and transfer cost data. The list of control variables includes: a time trend (the associated parameter to be estimated is denoted β_{time}), eleven monthly dummy variables and four daily variables to control for seasonal effects, and two dummy variables: $D_{2004-2005}$ that takes the value 1 during the period covering October 1, 2004, to September 30, 2005 (the parameter to be estimated is denoted $\beta_{D_{2004-2005}}$), and $D_{2005-2006}$ that takes the value 1 after October 1, 2005 (the parameter to be estimated is denoted $\beta_{D_{2005-2006}}$). Each of these period corresponds to a "standard gas year" during which the regulated Entry-Exit tariff system used by the UK National Transportation System is kept unchanged.

4.3 Descriptive statistics

An examination of trade data indicates that out of these 723 observations, 369 correspond to net positive exports to Belgium (of which 26 correspond to a congested infrastructure), 341 to net imports to the UK (of which 46 correspond to a congested infrastructure) and 13 to zero trade.

Table 2 summarizes the descriptive statistics for the two series R_{jit} (i.e., the difference between the spatial price spread and the unit transfer cost) both for the entire sample and for a restricted sample that omits the two exceptional episodes. The distributional properties of these series show some signs of non-normality as a very large leptokurtosis is observed in both cases. The estimated first-order autocorrelation coefficients reveal evidence of serial correlation. This finding is in favor of a dynamic specification able to correct for serial correlation.

Table 2. Descriptive statistics for the marginal rent to spatial arbitrage

	$R_{NBP \to ZEE}$	$R_{ZEE \to NBP}$
Mean	-0.232	0.044
Median	-0.100	-0.051
Maximum	7.484	25.189
Minimum	-25.543	-8.096
Standard Deviation	1.581	1.557
Skewness	-7.394	7.396
Kurtosis	108.946	112.664
K-S test	0.104	0.261
(p-value)	*(0.000)*	*(0.000)*
First-order autocorrelation	0.311^{***}	0.301^{***}
Observations	723	723

Note: K-S is the Kolmogorov-Smirnov test for the null hypothesis of normality. Asterisks indicate significance at the 0.01^{***} level.

Figure 1. Data plots

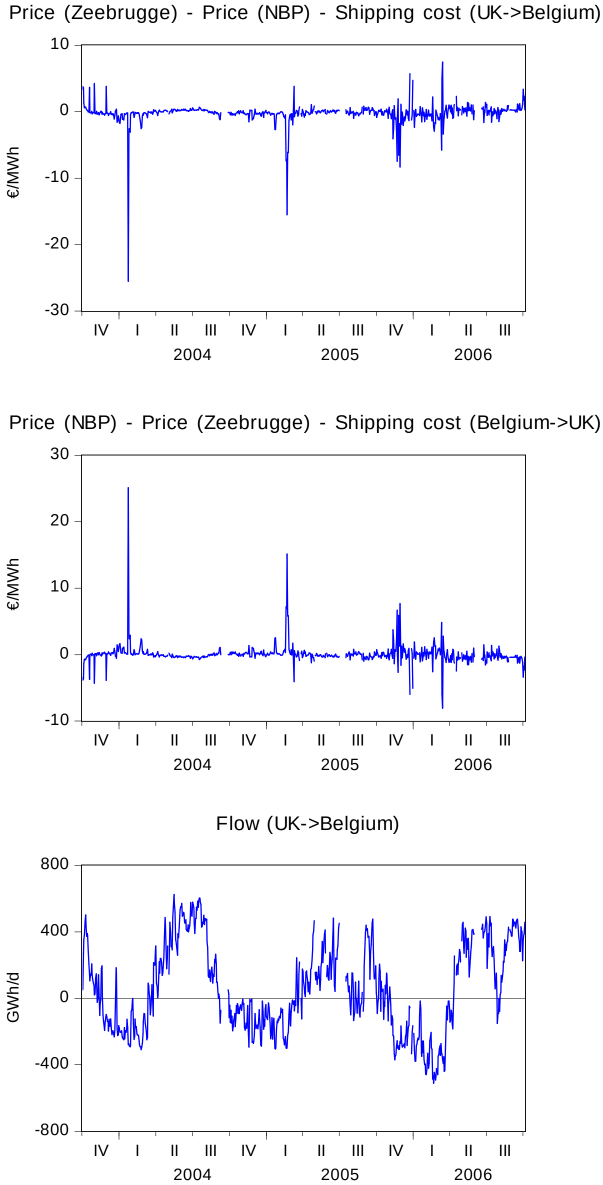

Figure 1 provides plots of these two series R_{jit} and the measured pipeline flow from Bacton (UK) to Zeebrugge (Belgium). A visual inspection of these plots suggests that the series R_{jit} exhibit two very large spikes on January 23, 2004 and between February 18 and February 24, 2005. According to market commentators, a conjunction of exceptional factors posed dramatic upward pressure on the UK NBP prices in these two occasions. On Friday 23 January 2004, an outage affected the withdrawal operations conducted at the Rough storage site during a particularly cold weather episode. As this storage site accounts for about 70% of the UK's gas storage capacity, this outage prompted a "dash for gas" that resulted in exuberantly high prices. On February 18th, 2005, a sudden and colder-than-anticipated weather episode began in the UK creating a need for immediate injections of LNG into the UK national transportation system at a moment when there was no available LNG cargoes that could had been redirected to the UK. During these spikes, the marginal rents to arbitrage from the UK to

Belgium (respectively from the continent to the UK) were obviously very negative (respectively positive).

Given their magnitude, one could wonder whether the presence of these spikes is consistent with the modeling assumptions retained for the truncated random variables μ_{jit} and υ_{jit}. In particular, it has to be verified whether the use of a time-invariant variance parameter σ_{υ}^{2} for the arbitrages performed from the UK to Belgium (σ_{μ}^{2} for the opposite direction) holds. To this purpose, we introduce $D_{23\text{ Jan }2004}$ and $D_{18-24\text{ Feb }2005}$, two dummy variables that take the value one during their respective episode and zero elsewhere, and allow the standard deviation of the half-normal random variable υ_{jit} (respectively μ_{jit}) to be of the form $\sigma_{\upsilon}+\zeta_{D_{23\text{ Jan }2004}}D_{23\text{ Jan }2004}+\zeta_{D_{18-24\text{ Feb }2005}}D_{18-24\text{ Feb }2005}$ (respectively $\sigma_{\mu}+\zeta_{D_{23\text{ Jan }2004}}D_{23\text{ Jan }2004}+\zeta_{D_{18-24\text{ Feb }2005}}D_{18-24\text{ Feb }2005}$) where $\zeta_{D_{23\text{ Jan }2004}}$ and $\zeta_{D_{18-24\text{ Feb }2005}}$ are parameters to be estimated, for the arbitrages performed from the UK to Belgium (respectively from Belgium to UK).

4.4 Estimation and empirical results

<u>a – Estimation procedure</u>

The estimation procedure involves the constrained maximization of a non-trivial log-likelihood function. This is a non-linear, non-convex, constrained optimization problem that has to be solved numerically using hill-climbing procedures.[10]

To obtain a feasible starting point, we first consider the simplest possible static specification (i.e., omitting the time trend, the dummy variables, the flow variable, the autocorrelation, the GARCH parameters and the spike parameters). The converged solution for this restricted specification is then used as a feasible starting point for the unrestricted models. The optimization problem at hand has the potential for local maxima, which is a source of concern because the outcome of a non-linear programming solver may depend on the location of the starting point. To address this problem, the first solution is compared to the ones obtained with a sample of 500 starting points uniformly drawn over a range of possible starting values. The converged solution that provides the highest likelihood value is systematically stored.

<u>b – Empirical results</u>

We first consider the simplest static specification of the PBM and successively estimate two versions of it: the simple one presented in Section 3 which is hereafter labeled Model I.a; and an extended version, labeled Model I.b, where the dummy variables corresponding to the observed spikes are introduced in the standard deviations of the respective truncated random variables (i.e., υ for the arbitrages performed from Belgium to the UK and μ in the opposite direction).

[10] *All the estimates reported in this paper have been obtained using an iterative procedure that performs 20 iterations using the Davidon-Fletcher-Powell (DFP) algorithm followed by 20 iterations using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) one, and then a switch back to DFP for 20 iterations, and so forth.*

The estimation results are presented in the first two columns of Table 3 for the arbitrages performed from the UK to Belgium, and of Table 4 for the arbitrages performed in the opposite direction. These tables detail the estimates obtained for: the market power coefficient (γ), the regime probabilities (λ's), the measurement bias parameters (α, β_{time}, $\beta_{D_{2004-2005}}$, $\beta_{D_{2005-2006}}$), the standard deviation parameters for the normal and truncated normal distributions (σ_{ε}, σ_{μ}, σ_{υ}, $\zeta_{D_{23\text{ Jan }2004}}$, $\zeta_{D_{18-24\text{ Feb }2005}}$) and a likelihood ratio test of the null hypothesis $\gamma = 0$. For concision, the seasonal parameters included in the mean equation (i.e., the coefficients of the daily and monthly dummy variables) are not reported.

From these estimation results, four main lines of findings can be highlighted. First, the estimated values for the coefficient γ are positive, as expected. These estimates are highly significant in both directions, which reveals the presence of imperfectly competitive arbitrages across the Channel. A further confirmation is provided by the likelihood ratio tests: the null hypothesis of competitive arbitrages is firmly rejected in both directions. So, we cannot reject the assumption of imperfectly competitive arbitrages during that period. This finding is consistent with the results in Rupérez-Micola and Bunn (2007).

Second, the high estimates obtained for λ_I and λ_{II} in both directions reveal that the observed spatial price difference is predominantly explained by the sum of the unobserved marginal transaction costs and the markup term. These very high values result in a very high probability of observing a spatial market equilibrium. Following Barrett and Li (2002), the probability of spatial market equilibrium conditions holding is in the range defined by the minimum and the maximum values of the direction-specific sums $\left(\lambda_I + \lambda_{II} + \lambda_{III_b} + \lambda_{VI}\right)$, that is (91.635, 92.340) for Model I.a and (89.085, 91.184) for Model I.b.

Third, the estimated probabilities λ_{III_a}, though small, are positive. Infrastructure congestion issues that are directly related to the Interconnector pipeline cannot be invoked to explain the presence of these strictly positive marginal profits to spatial arbitrage. These observed trade barriers could, for example, be due to pipeline congestion in the adjacent systems. In contrast, the probabilities λ_{III_b} to jointly observe infrastructure congestion and strictly positive marginal profits to spatial arbitrage regime are either zero or very low. These estimated values are consistent with the analysts' consensus summarized in Futyan (2006) on: (i) the oversized nature of the IUK's transportation capacity when natural gas is flowing to the Continent and (ii) the likely capacity-constrained nature of the IUK in the opposite direction (before the November 2005 capacity increase).

Lastly, one can compare the magnitude of the log-likelihoods of these two static models. In both directions, the null hypothesis $\zeta_{D_{23\text{ Jan }2004}} = \zeta_{D_{18-24\text{ Feb }2005}} = 0$ is firmly rejected by a standard likelihood ratio test which confirms the need to allow for a dedicated modeling of the spikes.

Table 3. Estimation results for natural gas trade from the UK to Belgium

	MODEL I (static)		**MODEL II (dynamic)**	
	Model I.a	Model I.b	Model II.a	Model II.b
Mean parameters				
α	-0.6429***	-0.5679***	-0.3173***	-0.2321*
β_{time}	5.0668***	4.9384***	1.5368	0.3255
$\beta_{D_{2004-2005}}$	-1.6190***	-1.5667***	-0.4331	-0.0412
$\beta_{D_{2005-2006}}$	-3.2620***	-3.1411***	-1.1614*	-0.3600
γ	0.0013***	0.0013***	0.0013***	0.0012***
ρ	______	______	0.3592***	0.4857***
Second moment parameters				
Regimes I & II				
σ_ε	0.2336***	0.2029***	______	______
ϖ	______	______	0.0186***	0.0118
δ	______	______	0.9227***	0.9406***
φ	______	______	0.0161	0.0120
Regimes III_a, III_b & IV				
σ_μ	2.3515***	2.2738***	2.1237***	1.8522***
Regimes V & VI				
σ_υ	3.7553***	1.6997***	6.2318***	1.5528***
$\zeta_{D_{23\text{ Jan }2004}}$	______	23.0378	______	27.3104***
$\zeta_{D_{18-24\text{ Feb }2005}}$	______	7.3792***	______	2.2908
Probabilities (in %)				
λ_I	47.2321***	46.0367***	48.0096***	46.0932***
λ_{II}	32.1778***	27.2187***	41.9937***	36.7620***
λ_{III_a}	2.9901***	3.3179***	3.1748***	4.4911***
λ_{III_b}	0.0000	0.0000	0.0000	0.0000
λ_{IV}	3.8551***	3.8155***	2.7430***	3.8501***
λ_V	0.8150**	1.6822***	0.0002	0.7116
λ_{VI}	12.9298***	17.9290***	4.0787***	8.0920***
Log likelihood	-1070.2260	-1006.8927	-967.7803	-934.2613
Akaike Information Criterion	2200.452	2077.7854	2001.5606	1938.5226
LR tests				
H_0: $\gamma = 0$	143.638 *(0.000)*	163.851 *(0.000)*	122.202 *(0.000)*	165.614 *(0.000)*
H_0: $\rho = \delta = \varphi = 0$	______	______	204.891 *(0.000)*	145.263 *(0.000)*
Observations	723	723	723	723

Note: Estimates for the monthly and daily dummies are not reported for brevity. Significance tests are based on asymptotic standard errors that have been computed using the Hessian matrix of the log-likelihood function. Asterisks indicate significance at 0.10*, 0.05** and 0.01*** levels, respectively. Numbers in parentheses are the *p*-values of the χ^2 statistics of the likelihood ratio (LR) tests.

Table 4. Estimation results for natural gas trade from Belgium to the UK

	MODEL I (static)		MODEL II (dynamic)	
	Model I.a	Model I.b	Model II.a	Model II.b
<u>Mean parameters</u>				
α	-0.2305***	-0.2513***	-0.0092	-0.0300
β_{time}	-0.4032	-0.2824	-3.2311***	-2.9068***
$\beta_{D_{2004-2005}}$	0.1674	0.1311	1.1062***	0.9939***
$\beta_{D_{2005-2006}}$	0.1057	0.0349	1.8571***	1.6357***
γ	0.0027***	0.0027***	0.0022***	0.0021***
ρ	______	______	0.3738***	0.3770***
<u>Second moment parameters</u>				
Regimes I & II				
σ_ε	0.2938***	0.2650***	______	______
ϖ	______	______	0.0199***	0.0178***
δ	______	______	0.9292***	0.9643***
φ	______	______	0.0042	0.0000
Regimes III_a, III_b & IV				
σ_μ	5.3064***	2.0717***	5.9213***	1.9007***
$\zeta_{D_{23\ Jan\ 2004}}$	______	22.5982	______	22.3679
$\zeta_{D_{18-24\ Feb\ 2005}}$	______	6.3602**	______	5.0316
Regimes V & VI				
σ_υ	2.5157***	2.3382***	2.3246***	2.1070***
<u>Probabilities (in %)</u>				
λ_I	37.8438***	35.6264***	41.6279***	40.9389***
λ_{II}	48.9237***	47.6156***	49.9429***	48.5970***
λ_{III_a}	2.8774***	4.3503***	1.9708***	2.7345***
λ_{III_b}	1.9321***	2.1687***	0.5832**	0.5764*
λ_{IV}	0.9762**	1.5448**	0.9646***	1.7067***
λ_V	4.5118***	5.0199***	2.7215***	2.8430***
λ_{VI}	2.9350***	3.6743***	2.1890***	2.6036***
Log likelihood	-1085.9403	-1057.9039	-956.6066	-925.5461
Akaike Information Criterion	2231.8806	2179.8078	1979.2132	1921.0922
LR tests				
H_0: $\gamma = 0$	146.538 *(0.000)*	134.329 *(0.000)*	112.599 *(0.000)*	96.105 *(0.000)*
H_0: $\rho = \delta = \varphi = 0$	______	______	258.667 *(0.000)*	264.716 *(0.000)*
Observations	723	723	723	723

Note: Estimates for the monthly and daily dummies are not reported for brevity. Significance tests are based on asymptotic standard errors that have been computed using the Hessian matrix of the log-likelihood function. Asterisks indicate significance at 0.10*, 0.05** and 0.01*** levels, respectively. Numbers in parentheses are the *p*-values of the χ^2 statistics of the likelihood ratio (LR) tests.

As such a static specification is *de facto* poorly adapted to capture the dynamic properties of daily data, we also consider the enriched dynamic specification presented in Appendix C that includes a correction for first-order serial correlation and a GARCH(1,1) effect. Again, two versions are successively considered. In Model II.a, the standard deviations of the half-normal random variables are time-invariant whereas Model II.b includes the dummy variables $D_{23\text{ Jan }2004}$ and $D_{18-24\text{ Feb }2005}$ to control for the effects of these exceptional episodes. The estimation results are presented in the third and fourth columns of Table 3 and Table 4. Compared to the static model, these two columns also report the estimates of the autocorrelation parameter (ρ), the GARCH parameters (ϖ, δ, φ) used to model the heteroscedasticity of the residuals in regimes I and II, and a likelihood ratio test of the null hypothesis of a static model (i.e., $\rho = \delta = \varphi = 0$).

From the estimation results, we observe that the estimated autocorrelation coefficients ρ and the estimated ARCH coefficients δ are highly significant in both directions which justifies the use of a dynamic specification. A confirmation is provided by the likelihood ratio tests that show that the null hypothesis of a static model (i.e., $\rho = \delta = \varphi = 0$) is firmly rejected for both Model II.a and Model II.b in both directions.

Regarding the interpretation, it should be noted that the estimation results obtained with the dynamic specification are consistent with the findings obtained with the static one. Again, the estimates document the imperfect nature of the competition among spatial arbitragers because the market power coefficient γ is positive and highly significant in both Model II.a and II.b in both directions. Moreover, the likelihood ratio tests again firmly reject the null hypothesis of competitive arbitrages in both directions. Regarding the regime probabilities, the estimates obtained for λ_I and λ_{II} remain the largest in both directions. The probability of spatial market equilibrium conditions holding is higher than 90% as the range defined by the minimum and the maximum values of the direction-specific sums $\left(\lambda_I + \lambda_{II} + \lambda_{III_b} + \lambda_{VI}\right)$ are (94.082, 94.343) for Model II.a and (90.947, 92.716) for Model II.b).

5. Concluding remarks

In Europe, the question of how to detect market power in the spatial arbitrages observed in a restructured natural gas industry is one of the key challenges that regulators and competition authorities have to address. The objective of this paper is to offer an empirical methodology which is able to test for the presence of perfect competition in these spatial arbitrages. Our approach explicitly builds upon the literature dedicated to natural gas markets integration and extends it by focusing on the relationship between the observed spatial price difference and the intermarket trade flows.

A case study focusing on the IUK pipeline during the period 2003–2006 provided us with an opportunity to obtain a series of original findings. The estimated probability of spatial market equilibrium conditions holding is very high, suggesting high degrees of wholesale natural gas market integration, consistent with previous research on IUK price co-movements (Neumann et al., 2006). But, the empirical evidence also suggests the presence of imperfect competition in the observed spatial arbitrages, consistent with the price-data results in Rupérez-Micola and Bunn (2007). Although our discussion is centered on this specific infrastructure, it should be clear these results imply that some care is needed when interpreting the high degree of co-movements which is typically documented in the empirical studies conducted on European spatial market price data. Though these co-movements can be interpreted as objective signs of market integration, they do not necessarily reveal the existence of a perfectly competitive internal market.

The institutional arrangements implemented in the UK to govern the functioning of the natural gas market have influenced the design of the other EU gas markets (Heather, 2010; Hallack and Vazquez, 2013). Future research will thus examine whether or not market equilibrium conditions hold in less mature continental markets. Such research could be useful for informing the current EU regulatory debates related to the functioning of the internal market for natural gas. From a methodological perspective, such research could also explore the possibility to opt for a more general class of model (e.g., a hidden Markov specification) that could, for example, allow for a possibly changing behaviour of the spatial arbitrageurs (e.g., to represent agents that may adopt a competitive behaviour in some observations and exert market power in others).

References

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

Abada, I., Gabriel, S.A., Briat, V., Massol, O. (2013). "A Generalized Nash–Cournot Model for the Northwestern European Natural Gas Markets with a Fuel Substitution Demand Function: The GaMMES Model." *Networks and Spatial Economics* 13(1): 1–42.

Argentesi, E., Banal-Estañol, A., Seldeslachts, J., Andrews, M. (2017) "A retrospective evaluation of the GDF/Suez merger: Effects on gas hub prices." *DIW Discussion Paper 1664, DIW Berlin*.

Asche, F., Osmundsen, P., Tveterås, R. (2002). "European Market Integration for Gas? Volume Flexibility and Political Risk." *Energy Economics* 24(3): 249–265.

Asche, F., Misund, B., Sikveland, M. (2013). "The relationship between spot and contract gas prices in Europe." *Energy Economics* 38: 212–217.

Bailey, E.M. (1998). *Empirical Studies of New Markets in the US Electricity Industry*. Department of Economics, Massachusetts Institute of Technology, unpublished Ph.D. thesis.

Barrett, C.B. (1996). "Market Analysis Methods: Are Our Enriched Toolkits Well Suited to Enlivened Markets?" *American Journal of Agricultural Economics* 78(3): 825–829.

Barrett, C.B. (2001). “Measuring Integration and Efficiency in International Agricultural Markets.” *Review of Agricultural Economics* 23(1): 19–32.

Barrett, C.B., Li, J.R. (2002). “Distinguishing between Equilibrium and Integration in Spatial Price Analysis.” *American Journal of Agricultural Economics* 84(2): 292–307.

Baulch, B. (1997). “Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration.” *American Journal of Agricultural Economics* 79(2): 477–487.

Borenstein, S., Bushnell, J., Knittel, C. R., Wolfram, C. (2008). “Inefficiencies and market power in financial arbitrage: a study of California's electricity markets.” *The Journal of Industrial Economics*, 56(2), 347-378.

Bollerslev, T. (1986). “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics* 31(3): 307–327.

Brown, S.P.A., Yücel, M.K. (2008). “Deliverability and regional pricing in U.S. natural gas markets.” *Energy Economics* 30(5): 2441–2453.

Cirera, X., Arndt, C. (2008), “Measuring the impact of road rehabilitation on spatial market efficiency in maize markets in Mozambique.” *Agricultural Economics*, 39: 17–28.

Cottle, R.W., Pang, J.S., Stone, R.E. (2009). “The linear complementarity problem.” Society for Industrial Mathematics: Philadelphia.

Cuddington, J.T., Wang, Z. (2006). “Assessing the degree of spot market integration for US natural gas: evidence from daily price data.” *Journal of Regulatory Economics* 29(2): 195–210.

Day, C.J, Bunn, D.W. (2001). “Divestiture of Generation Assets in the Electricity Pool of England and Wales: A Computational Approach to Analyzing Market Power.” *Journal of Regulatory Economics* 19(2): 123–141.

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

Deaton, A., Laroque, G. (1996). “Competitive Storage and Commodity Price Dynamics.” *Journal of Political Economy* 104(5): 896–923.

Doane, M.J., Spulber, D.F. (1994). “Open Access and the Evolution of the U. S. Spot Market for Natural Gas.” *Journal of Law and Economics* 37(2): 477–517.

Enke S. (1951). “Equilibrium among Spatially Separated Markets: Solution by Electric Analogue.” *Econometrica* 19(1): 40–47.

Facchinei, F., Kanzow, C. (2010). “Generalized Nash Equilibrium Problems.” *Annals of Operations Research* 175(1): 177–211.

Fackler, P., Goodwin, B. (2001). “Spatial Price Analysis.” In *Handbook of Agricultural Economics*, ed. B. L. Gardner and G. C. Rausser, 971–1024. Amsterdam: Elsevier Science.

Futyan, M. (2006). *The Interconnector Pipeline: a Key Link in Europe’s Gas Network.* NG11. Oxford Institute of Energy Studies.

Gabriel, S.A., Conejo, A.J., Hobbs, B.F., Fuller, D., Ruiz, C. (2013) “Complementarity Modeling In Energy Markets.” Springer: New York.

Golombek, R, Gjelsvik, E, Rosendahl, KE. (1995). “Effects of liberalizing the natural gas markets in Western Europe.” *The Energy Journal* 16(1): 85–111.

Green, R. (1999). “The Electricity Contract Market in England and Wales.” *The Journal of Industrial Economics* 47(1): 107–124.

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.

Harker, P.T., (1986). “Alternative Models of Spatial Competition.” *Operations Research* 34(3): 410–425.

Heather, P., (2010). *The evolution and functioning of the traded gas market in Britain*. NG44. Oxford Institute of Energy Studies.

Hubbard, R. G., Weiner, R. J. (1986). “Regulation and long-term contracting in US natural gas markets.” *The Journal of Industrial Economics* 35(1): 71–79.

Huntington, H.G. (2009). “World Natural Gas Markets And Trade: A Multi-Modeling Perspective.” *The Energy Journal*, Special Issue.

King, M., Cuc M. (1996). “Price Convergence in North American Natural Gas Spot Markets.” *The Energy Journal* 17(2): 17–42.

Kiefer, N.M. (1980). “A Note on Switching Regressions and Logistic Discrimination.” *Econometrica* 48(4): 1065–1069.

Kleit, A.N. (1998). “Did Open Access integrate Natural Gas Markets? An Arbitrage Cost Approach.” *Journal of Regulatory Economics* 14(2): 19–33.

Kleit, A.N. (2001). “Are Regional Oil Markets Growing Closer Together? An Arbitrage Cost Approach.” *The Energy Journal* 22(2): 1–15.

Massol, O. (2011). “A Cost Function for the Natural Gas Transmission Industry: Further Considerations.” *The Engineering Economist* 56(2): 95–122.

McNew, K., Fackler, P.L. (1997). "Testing Market Equilibrium: Is Cointegration Informative?" *Journal of Agricultural and Resource Economics* 22(2): 191–207.

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

Moser, C., Barrett, C., Minten, B. (2009). "Spatial Integration at Multiple Scales: Rice Markets in Madagascar." *Agricultural Economics* 40(3): 281–294.

Negassa, A., Myers, R.J. (2007). "Estimating Policy Effects on Spatial Market Efficiency: An Extension to the Parity Bounds Model." *American Journal of Agricultural Economics* 89(2): 338–352.

Neumann, A., Siliverstovs, B., von Hirschhausen, C. (2006). "Convergence of European spot market prices for natural gas? A real-time analysis of market integration using the Kalman Filter." *Applied Economics Letters* 13(11): 727–732.

Olsen, K.K., Mjelde, J.W., Bessler, D.A. (2015). "Price formulation and the law of one price in internationally linked markets: an examination of the natural gas markets in the USA and Canada." *The Annals of Regional Science* 54(1) 117–142.

Park, H., Mjelde, J.W., Bessler, D.A. (2008). "Price interactions and discovery among natural gas spot markets in North America." *Energy Policy* 36(1): 290–302.

Platt's (2012). *European Natural Gas Assessments and Indices. Methodology and Specification Guide. November 2012*.

Renou-Maissant, P. (2012). "Toward the Integration of European Natural Gas Markets: A time-varying Approach." *Energy Policy* 51: 779–790.

Ritz, R.A., (2014) "Price discrimination and limits to arbitrage: An analysis of global LNG markets." *Energy Economics*, 45: 324-332.

Ruiz, C., Conejo, A.J., Fuller, J.D., Gabriel, S.A., Hobbs, B.F. (2014). "A Tutorial Review of Complementarity Models for Decision-Making in Energy Markets," *EURO Journal of Decision Processes*, 2(1-2), 91-120.

Rupérez Micola, A., Bunn, D.W. (2007). "Two Markets and a Weak Link." *Energy Economics* 29(1): 79–93.

Samuelson, P.A. (1952). "Spatial Price Equilibrium and Linear Programming." *The American Economic Review* 42(3): 283–303.

Serletis, A. (1997). "Is There an East-West Split in North American Natural Gas Markets?" *The Energy Journal* 18(1): 47–62.

Sexton, R., Kling, C., Carman, H. (1991). "Market Integration, Efficiency of Arbitrage and Imperfect Competition: Methodology and Application to U.S. Celery." *American Journal of Agricultural Economics* 73(3): 568–580.

Siliverstovs, B., L'Hégaret, G., Neumann, A., von Hirschhausen, C. (2005). "International market integration for natural gas? A cointegration analysis of prices in Europe, North America and Japan." *Energy Economics* 27(4): 603–615.

Spiller, P.T., Huang, C.J. (1986). "On the Extent of the Market: Wholesale Gasoline in the Northern United States." *Journal of Industrial Economics* 35(2): 131–145.

Spiller, P.T., Wood, R.O. (1988). "Arbitrage during the Dollar-Sterling Gold Standard, 1899-1908: An Econometric Approach." *Journal of Political Economy* 96(4): 882–892.

Stigler, G.J., Sherwin, R.A. (1985). "The Extent of the Market." *Journal of Law and Economics* 28(3): 555–585.

Takayama, T., Judge, G.C. (1971). "Spatial and Temporal Price and Allocation Models." North-Holland: Amsterdam.

Vazquez, M., Hallack, M., Glachant, J-M. (2012). "Designing the European Gas Market: More Liquid & Less Natural?" *Economics of Energy & Environmental Policy* 1(3): 25–38.

Weinstein, M.A. (1964). "The Sum of Values from a Normal and a Truncated Normal Distribution." *Technometrics* 6: 104–5.

Wright, P. (2006). "Gas Prices in the UK. Markets and Insecurity of Supply." Oxford Institute for Energy Studies. Oxford University Press: Oxford.

Yépez, R.A. (2008). "A cost function for the natural gas industry." *The Engineering Economist*, 53(1), 68–83.

Zant, W. (2013). "How Is the Liberalization of Food Markets Progressing? Market Integration and Transaction Costs in Subsistence Economies." *The World Bank Economic Review*, 27(1): 28–54.

Appendix A – Perfectly competitive spatial arbitrages

This Appendix details the formal proof of Proposition 1. We let: D_{it} and D_{jt} denote the quantities demanded at time t in each region, $B_i(D_{it}) = \int_0^{D_{it}} p_{it}^D(q)\,dq$ and $B_j(D_{jt}) = \int_0^{D_{jt}} p_{jt}^D(q)\,dq$ denote the gross consumer surplus in each region at that time, and $C_i(S_{it}) = \int_0^{S_{it}} p_{it}^S(q)\,dq$ and $C_j(S_{jt}) = \int_0^{S_{jt}} p_{jt}^S(q)\,dq$ denote the total cost incurred by the producers in each region. By construction, we have: $B_i^{'}(D_{it}) = p_{it}^D(D_{it})$; $B_j^{'}(D_{jt}) = p_{jt}^D(D_{jt})$; $C_i^{'}(S_{it}) = p_{it}^S(S_{it})$ and $C_j^{'}(S_{jt}) = p_{jt}^S(S_{jt})$.

Proof of Proposition 1: *Following Takayama and Judge (1971), we define the net quasi-welfare function for the two regions as:*

$$W\left(D,S,Q_{jit}\right)=B_i\left(D_{it}\right)+B_j\left(D_{jt}\right)-C_i\left(S_{it}\right)-C_j\left(S_{jt}\right)-T_{jit}Q_{jit}\,, \tag{A.1}$$

this form is also known as the net social payoff which maximizes the sum of producers' and consumers' surplus after deducting for transfer costs (Samuelson, 1952). The competitive market equilibrium is the solution of the following optimization problem (Takayama and Judge, 1971):

$$\underset{D,S,Q_{jit}}{Max} \quad W\left(D,S,Q_{jit}\right) \tag{A.2}$$

$$s.t. \quad D_{it}-S_{it}-Q_{jit}\le 0 \tag{A.3}$$

$$D_{jt}+Q_{jit}-S_{jt}\le 0 \tag{A.4}$$

$$Q_{jit}\le K_{jit} \tag{A.5}$$

$$D_{it}\ge 0\,,\ D_{jt}\ge 0\,,\ S_{it}\ge 0\,, S_{jt}\ge 0\,,\ Q_{jit}\ge 0\,. \tag{A.6}$$

where the objective is to maximize the net social payoff subject to a set of linear constraints. Here, the constraint (A.3) states that the quantity consumed in market i is less than or equal to the sum of the local supply and the quantity shipped into that region; the constraint (A.4) states that the sum of the quantity consumed in market j and the quantity shipped from that region is less than or equal to the local supply; and (A.5) is the capacity constraint related to the transportation infrastructure. We let: ϕ_{it}, ϕ_{jt} and ξ_{jit} denote the dual variables associated with the constraints (A.3), (A.4) and (A.5) respectively. The dual variables ϕ_{it} and ϕ_{jt} can be interpreted as the market clearing prices in the two markets and the dual variable ξ_{jit} can be interpreted as the market clearing price in the market for transportation rights. Hereafter, we let L denote the Lagrangian for this optimization problem.

As the objective function is quadratic and strictly concave with respect to both D and S, this problem has a unique solution $\left(D_{it}^*,D_{jt}^*,S_{it}^*,S_{jt}^*,Q_{jit}^*,\phi_{it}^*,\phi_{jt}^*,\xi_{jit}^*\right)$ that verifies the Karush-Kuhn-Tucker optimality conditions listed in Table A.1.

Table A.1. The Karush-Kuhn-Tucker optimality conditions for the optimization problem

$\dfrac{\partial L}{\partial D_{it}}$:	$p_{it}^{D}\left(D_{it}\right)-\phi_{it}\le 0$, $\quad D_{it}\ge 0$ $\quad$ and $\quad \left[p_{it}^{D}\left(D_{it}\right)-\phi_{it}\right]D_{it}=0$,	(A.7)
$\dfrac{\partial L}{\partial D_{jt}}$:	$p_{jt}^{D}\left(D_{jt}\right)-\phi_{jt}\le 0$, $\quad D_{jt}\ge 0$ $\quad$ and $\quad \left[p_{jt}^{D}\left(D_{jt}\right)-\phi_{jt}\right]D_{jt}=0$,	(A.8)
$\dfrac{\partial L}{\partial S_{it}}$:	$-p_{it}^{S}\left(S_{it}\right)+\phi_{it}\le 0$, $\quad S_{it}\ge 0$ $\quad$ and $\quad \left[-p_{it}^{S}\left(S_{it}\right)+\phi_{it}\right]S_{it}=0$,	(A.9)
$\dfrac{\partial L}{\partial S_{jt}}$:	$-p_{jt}^{S}\left(S_{jt}\right)+\phi_{jt}\le 0$, $\quad S_{jt}\ge 0$ $\quad$ and $\quad \left[-p_{jt}^{S}\left(S_{jt}\right)+\phi_{jt}\right]S_{jt}=0$,	(A.10)
$\dfrac{\partial L}{\partial Q_{jit}}$:	$-T_{jit}+\phi_{it}-\phi_{jt}-\xi_{jit}\le 0$, $\quad Q_{jit}\ge 0$ $\quad$ and $\quad \left[-T_{jit}+\phi_{it}-\phi_{jt}-\xi_{jit}\right]Q_{jit}=0$,	(A.11)
$\dfrac{\partial L}{\partial \phi_{it}}$:	$D_{it}-S_{it}-Q_{jit}\le 0$, $\quad \phi_{it}\ge 0$ $\quad$ and $\quad \left[D_{it}-S_{it}-Q_{jit}\right]\phi_{it}=0$,	(A.12)
$\dfrac{\partial L}{\partial \phi_{jt}}$:	$D_{jt}+Q_{jit}-S_{jt}\le 0$, $\quad \phi_{jt}\ge 0$ $\quad$ and $\quad \left[D_{jt}+Q_{jit}-S_{jt}\right]\phi_{jt}=0$,	(A.13)
$\dfrac{\partial L}{\partial \xi_{jit}}$:	$Q_{jit}\le K_{jit}$, $\quad \xi_{jit}\ge 0$ $\quad$ and $\quad \left[Q_{jit}-K_{jit}\right]\xi_{jit}=0$.	(A.14)

First, we examine the equilibrium at market i. *We reason by contradiction and assume that, at market equilibrium, the quantity supplied by the local producers is zero: i.e.,* $S_{it}^{*}=0$. *Condition (A.9) imposes* $\phi_{it}^{*}\le p_{it}^{S}\left(0\right)$ *i.e.,* $\phi_{it}^{*}\le c_{it}$. *Condition (A.7) indicates that* $p_{it}^{D}\left(D_{it}^{*}\right)\le\phi_{it}^{*}$ *and thus* $a_{it}-b_i\,D_{it}^{*}\le c_{it}$ *must hold. As* $S_{it}^{*}=0$, *condition (A.12) indicates that* $D_{it}^{*}\le Q_{jit}^{*}$ *and condition (A.14) indicates that* $Q_{jit}^{*}\le K_{jit}$. *As* $b_i>0$, *we must have* $a_{it}-b_i\,K_{jit}\le c_{it}$ *which contradicts the assumption* $a_{it}-b_i\,K_{jit}>c_{it}$ *(cf., Section 2) and thus invalidates the assumption* $S_{it}^{*}=0$. *Hence, the equilibrium is such that the local supply is positive* $S_{it}^{*}>0$ *which proves (cf. condition (A.9)) that the market price* ϕ_{it}^{*} *at the destination market is equal to the regional supply price, i.e.* $\phi_{it}^{*}=p_{it}^{S}\left(S_{it}^{*}\right)$. *As* $\phi_{it}^{*}>0$, *the condition (A.12) is such that* $D_{it}^{*}=S_{it}^{*}+Q_{jit}^{*}$. *As* $S_{it}^{*}>0$ *and* $Q_{jit}^{*}\ge 0$ *(cf., (A.11)), the quantity demanded at market* i *is positive:* $D_{it}^{*}>0$ *and the condition (A.7) reveals that, at market* i, *the market price* ϕ_{it}^{*} *verifies* $\phi_{it}^{*}=p_{it}^{D}\left(D_{it}^{*}\right)$. *Hence, the market price* ϕ_{it}^{*} *is equal to both* $p_{it}^{D}\left(D_{it}^{*}\right)$ *the price demanded by the local consumers and* $p_{it}^{S}\left(S_{it}^{*}\right)$ *the regional supply price. We let* P_{it} *denote that market clearing price.*

Second, we examine the equilibrium at market j. *Again, we reason by contradiction and assume that, at market equilibrium, the quantity demanded by the local consumers is zero: i.e.,* $D_{jt}^{*}=0$. *Condition (A.8) imposes* $p_{jt}^{D}\left(0\right)\le\phi_{jt}^{*}$ *that is:* $\phi_{jt}^{*}\ge a_{jt}$. *As* $\phi_{jt}^{*}>0$, *condition (A.13) indicates that* $S_{jt}^{*}=Q_{jit}^{*}$ *and thus* $p_{jt}^{S}\left(S_{jt}^{*}\right)=p_{jt}^{S}\left(Q_{jit}^{*}\right)$. *As* $Q_{jit}^{*}\le K_{jit}$, *we have* $p_{jt}^{S}\left(S_{jt}^{*}\right)\le p_{jt}^{S}\left(K_{jit}\right)$. *From condition*

(A.10), we thus have $\phi_{jt}^{*} \le p_{jt}^{S}\left(K_{jit}\right)$ which indicates that the condition $a_{jt} \le c_{it} + d_{i} K_{jit}$ must hold which contradicts the assumption $a_{jt} > c_{jt} + d_{j} K_{jit}$ (cf., Section 2). So, the equilibrium is such that the local demand must be positive $D_{jt}^{*} > 0$ which suggests that the market price ϕ_{jt}^{*} at the destination market is equal to $\phi_{jt}^{*} = p_{jt}^{D}\left(D_{jt}^{*}\right)$ the price demanded there (cf., condition (A.8)). As $D_{jt}^{*} > 0$ and $Q_{jit}^{*} \ge 0$, the condition (A.13) reveal that $D_{jt}^{*} + Q_{jit}^{*} \le S_{jt}^{*}$ which indicates that $S_{jt}^{*} > 0$ and thus, using condition (A.10), we obtain $\phi_{jt}^{*} = p_{jt}^{S}\left(S_{jt}^{*}\right)$. So, at market j, the market price ϕ_{jt}^{*} is equal to both $p_{ij}^{D}\left(D_{jt}^{*}\right)$ the price demanded by the local consumers and $p_{it}^{S}\left(S_{it}^{*}\right)$ the regional supply price. We let P_{jt} denote that market clearing price.

Substituting P_{jt} and P_{jt} for ϕ_{it}^{*} and ϕ_{jt}^{*} in condition (A.11), one can readily identify the complementarity conditions (1) and (2) in the conditions (A.11) and (A.14) that must hold at equilibrium. Q.E.D.

Appendix B – Oligopolistic spatial arbitrages

In this Appendix, we present the technical developments needed to prove Proposition 2. We examine the collective behavior of the G gas traders. We proceed as in Gabriel et al. (2013) and define a series of conditions that together characterize the spatial equilibrium at time t: first, the collection of G mathematical programming problems describing the traders' individual behavior, and second, a complementarity condition similar to (2) that ties the traders' individual optimization problems and describes the market clearing condition in the market for transportation rights.[11]

First, each trader $g \in \{1,\ldots,G\}$ is a profit maximizing agent that solves the following optimization problem:

$$\underset{q_{jit}^{g}}{Max} \quad \left[p_{it}^{D}\left(S_{it}\left(q_{jit}^{g}+q_{jit}^{-g}\right)+q_{jit}^{g}+q_{jit}^{-g}\right)-p_{jt}^{D}\left(S_{jt}\left(q_{jit}^{g}+q_{jit}^{-g}\right)-q_{jit}^{g}-q_{jit}^{-g}\right)-T_{jit}-\xi_{jit}\right]q_{jit}^{g} \tag{B.1}$$

$$\text{s.t.} \quad q_{jit}^{g} \ge 0$$

where the non-negative decision variable q_{jit}^{g} is the flow traded by g from market j to market i at time t, q_{jit}^{-g} is a short notation for the sum of the flows decided by the other traders, $p_{it}^{D}(.)$ and $p_{jt}^{D}(.)$ are the local inverse demand functions. The objective function (B.1) describes the total profit obtained by trader g. In that objective function, the unit revenue obtained in the destination market i is $p_{it}^{D}\left(S_{it}\left(q_{jit}^{g}+q_{jit}^{-g}\right)+q_{jit}^{g}+q_{jit}^{-g}\right)$. The total unit cost is the sum of: $p_{jt}^{D}\left(S_{jt}\left(q_{jit}^{g}+q_{jit}^{-g}\right)-q_{jit}^{g}-q_{jit}^{-g}\right)$ the

[11] From a technical perspective, this problem is an instance of what is known in the operations research community as a Generalized Nash equilibrium problem (which is also named a social equilibrium problem in economics). We refer to the survey in Facchinei ·and Kanzow (2010) for a comprehensive presentation of that type of problems and to Ruiz et al. (2014) for an overview of applications in the context of energy markets.

purchasing price in j, T_{jit} the unit transfer cost, and ξ_{jit} the price of a transportation right. The following reasoning is useful to further define the relations $S_{it}(.)$ and $S_{jt}(.)$ that characterize how, in each market, the local industries' aggregate supply varies as a function of the traders' decisions. Using $Q_{jit} = \sum_{g=1}^{G} q_{jit}^{g}$ as a short notation for the aggregate flow of gas transferred by all the traders, we can remark that: at market equilibrium at time t, the local demanded price at the destination market i (respectively the origin market j) is $p_{it}^{D}\left(S_{it}+Q_{jit}\right)$ (respectively $p_{jt}^{D}\left(S_{jt}-Q_{jit}\right)$) and that price must be equal to the price obtained with the local inverse supply function $p_{it}^{S}\left(S_{it}\right)$ (respectively $p_{jt}^{S}\left(S_{jt}\right)$). Therefore, the functions $S_{it}(.)$ and $S_{jt}(.)$ are:[12]

$$S_{it}\left(Q_{jit}\right)=\frac{a_{it}-c_{it}}{b_i+d_i}-\frac{b_i}{b_i+d_i}Q_{jit}, \quad \text{and} \quad S_{jt}\left(Q_{jit}\right)=\frac{a_{jt}-c_{jt}}{b_j+d_j}+\frac{b_j}{b_j+d_j}Q_{jit}. \tag{B.2}$$

Second, the market clearing condition at the market for transportation rights is given by the following complementarity condition that ensures that the price of a transportation right ξ_{jit} is equal to zero whenever $\sum_{g=1}^{G} q_{jit}^{g}$ the aggregate demand for transportation rights is lower that the supply (i.e., whenever the transportation capacity constraint is slack) and that ξ_{jit} is positive when supply falls short of demand (i.e, when this constraint is binding):

$$0 \le \xi_{jit}, \qquad \sum_{g=1}^{G} q_{jit}^{g} \le K_{jit} \qquad \text{and} \qquad \left(\sum_{g=1}^{G} q_{jit}^{g}-K_{jit}\right)\xi_{jit}=0, \tag{B.3}$$

Replacing the relations (B.2) in the objective functions (B.1), deriving the Karush-Kuhn-Tucker conditions of the traders' optimization problems and using P_{it} and P_{jt} as a short notations for the local prices (i.e.: $P_{it}=p_{it}^{D}\left(S_{it}\left(Q_{jit}\right)+Q_{jit}\right)$ and $P_{jt}=p_{jt}^{D}\left(S_{jt}\left(Q_{jit}\right)-Q_{jit}\right)$), we obtain the following set of complementarity conditions that together characterize the equilibrium conditions for oligopolistic spatial arbitrages at time t:

$$\begin{cases} 0 \le q_{jit}^{g}, \qquad P_{it}-P_{jt}-T_{jit}-\left(\dfrac{d_i b_i}{b_i+d_i}+\dfrac{d_j b_j}{b_j+d_j}\right)q_{jit}^{g}-\xi_{jit} \le 0 \qquad \text{and} \\ \left(P_{it}-P_{jt}-T_{jit}-\left(\dfrac{d_i b_i}{b_i+d_i}+\dfrac{d_j b_j}{b_j+d_j}\right)q_{jit}^{g}-\xi_{jit}\right)q_{jit}^{g}=0, \qquad \forall g \in \{1,\dots,G\}, \end{cases} \tag{B.4}$$

[12] *Recall that we assume that the condition* $a_{it}-b_i K_{jit} > c_{it}$ *is verified. As* $Q_{jit} \le K_{jit}$*, the local supply in the destination market* i *in equation* (B.2) *is positive. Similarly, we also assume that* $a_{jt} > c_{jt}+d_j K_{jit}$*. Hence, the local supply in the origin market* j *in equation* (B.2) *is such that the amount consumed in the origin market (i.e., the difference between local supplies and exports* $S_{jt}\left(Q_{jit}\right)-Q_{jit}$*) is positive.*

$$0 \le \xi_{jit}\,, \qquad \sum_{g=1}^{G} q_{jit}^{g} \le K_{jit} \qquad \text{and} \qquad \left(\sum_{g=1}^{G} q_{jit}^{g} - K_{jit}\right)\xi_{jit} = 0\,, \tag{B.5}$$

The economic interpretation of these conditions is similar to those detailed for the case of competitive arbitrages except that the traders' marginal profits to spatial arbitrage are now modified to account for the players' ability to exert market power in both markets. The marginal revenue obtained by trader g in the destination market is $P_{it} - \frac{d_i\, b_i}{b_i + d_i} q_{jit}^{g}$ where $-\frac{d_i\, b_i}{b_i + d_i} q_{jit}^{g}$ represents the marginal loss of revenue from getting a lower price for each of the units the player is selling there. The trader's marginal cost includes three distinct components: (i) the marginal purchase cost in the origin market $P_{jt} + \frac{d_j b_j}{b_j + d_j} q_{jit}^{g}$ where $\frac{d_j b_j}{b_j + d_j} q_{jit}^{g}$ represents the marginal cost increase from getting a higher price for each of the units the player is purchasing (i.e., the effect of the players' oligopsonistic behavior in market j); (ii) the marginal transfer cost T_{jit}, and (iii) the price of a transportation right ξ_{jit}.

> ***Lemma:*** *There exists a unique vector of individual decisions* q_{jit}^{g*} *and a unique price* ξ_{jit}^{*} *that verifies the equilibrium conditions* (B.4) *and* (B.5).

Proof: *The conditions* (B.4) *and* (B.5) *together define the linear complementarity problem LCP (m,M):* $z \ge 0$, $Mz + m \ge 0$ *and* $z^T\left(Mz + m\right) = 0$, *where:* $z = \left[q_{jit}^{1} \;\; \cdots \;\; q_{jit}^{G} \;\; \xi_{jit}\right]^T$; $M = \begin{bmatrix} A & B^T \\ -B & 0 \end{bmatrix}$ *is a real matrix where* B *is the all-ones row matrix of size* $1 \times G$ *and* $A = \left(\frac{d_i\, b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j}\right)\left(I_G + J_G\right)$ *where* I_G *the identity matrix of size* G *and* J_G *the all-ones square matrix of size* G; *and* $m = \left[\left(a_{jt} - b_j\,\frac{a_{jt} - c_{jt}}{b_j + d_j} - a_{it} + b_i\,\frac{a_{it} - c_{it}}{b_i + d_i}\right) \;\; \cdots \;\; \left(a_{jt} - b_j\,\frac{a_{jt} - c_{jt}}{b_j + d_j} - a_{it} + b_i\,\frac{a_{it} - c_{it}}{b_i + d_i}\right) \;\; K_{jit}\right]^T$. *Let* z *be a non-zero column vector of* $\left(G+1\right)$ *real numbers, we have* $z^T M z > 0$ *because* $z^T M z = \left(\frac{d_i\, b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j}\right)\left(2\sum_{\substack{i=1 \\ j \ge i}}^{G+1} z_i z_j\right)$ *and* $2\sum_{i=1}^{G+1}\left(z_i\right)^2 + 2\sum_{\substack{i,j \\ j>i}} z_i z_j = \sum_{i=1}^{G+1}\left(z_i\right)^2 + \left(\sum_{i=1}^{G+1} z_i\right)^2$. *Hence,* M *is positive definite. The Theorem 3.1.6 in Cottle et al., (2009, p. 141) indicates that, if* M *is positive definite, there exists a unique solution to the LCP (m,M).* *Q.E.D.*

From an empirical perspective, the experience gained with restructured natural gas markets indicates that the individual decisions q_{jit}^{g} are seldom publicly available whereas the aggregate trade flow Q_{jit} are. As in the case of competitive arbitrage, one may wonder whether there exists aggregate complementarity conditions that must hold in case of an equilibrium. This is precisely the aim of the following proposition which is presented in Section 2.b.

Proposition 2: *If the following aggregate complementarity conditions:*

$$\begin{cases} 0 \le Q_{jit}, \qquad P_{it} - P_{jt} - T_{jit} - \left(\dfrac{d_i b_i}{b_i + d_i} + \dfrac{d_j b_j}{b_j + d_j} \right) \dfrac{Q_{jit}}{G} - \xi_{jit} \le 0 \qquad \textit{and} \\ \left(P_{it} - P_{jt} - T_{jit} - \left(\dfrac{d_i b_i}{b_i + d_i} + \dfrac{d_j b_j}{b_j + d_j} \right) \dfrac{Q_{jit}}{G} - \xi_{jit} \right) Q_{jit} = 0, \end{cases} \tag{B.6}$$

$$0 \le \xi_{jit}, \qquad Q_{jit} \le K_{jit} \qquad \textit{and} \qquad \left(Q_{jit} - K_{jit} \right) \xi_{jit} = 0. \tag{B.7}$$

where P_{it} and P_{jt} are the local market clearing prices, do not hold at time t, the equilibrium conditions (B.4) *and* (B.5) *for oligopolistic spatial arbitrages are not verified at that time.*

Proof: *We reason by contradiction and assume that the equilibrium conditions* (B.4) *and* (B.5) *are verified and, using the lemma above, we let $z^* = \left(q_{jit}^{1\,*}, ..., q_{jit}^{G\,*}, \xi_{jit}^* \right)$ denote the unique vector of decision variables that verifies these conditions. To begin with, we are going to prove that z^* verifies: either $q_{jit}^{g\,*} = 0$ for any $g \in \{1, ..., G\}$, or $q_{jit}^{g\,*} > 0$ for any $g \in \{1, ..., G\}$. Let us assume that z^* is such that there jointly exists at least one trader g with $q_{jit}^{g\,*} > 0$ and at least one trader g' with $q_{jit}^{g'*} = 0$. According to the equilibrium conditions* (B.4)*, the marginal profits of these players are:*

$$\textit{Trader } g: \qquad P_{it}^* - P_{jt}^* - T_{jit} - \left(\frac{d_i b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j} \right) q_{jit}^{g\,*} - \xi_{jit}^* = 0, \tag{B.8}$$

$$\textit{Trader } g': \qquad P_{it}^* - P_{jt}^* - T_{jit} - \xi_{jit}^* \le 0, \tag{B.9}$$

where $P_{it}^* = p_{it}^D \left(S_{it} \left(Q_{jit}^* \right) + Q_{jit}^* \right)$ and $P_{jt}^* = p_{jt}^D \left(S_{jt} \left(Q_{jit}^* \right) - Q_{jit}^* \right)$. Subtracting (B.8) from (B.9), we obtain $\left(\frac{d_i b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j} \right) q_{jit}^{g\,*} \le 0$ which is a contradiction because we have assumed that $q_{jit}^{g\,*}$, d_i, d_j, b_i, and b_j are all positive numbers. Hence, z^* has to be such that: either $q_{jit}^{g\,*} = 0$ for any $g \in \{1, ..., G\}$, or $q_{jit}^{g\,*} > 0$ for any $g \in \{1, ..., G\}$.

Using a similar argument, we can also prove that, if the equilibrium is such that $q_{jit}^{g\,*} > 0$ for any $g \in \{1, ..., G\}$, we must have $q_{jit}^{g\,*} = Q_{jit}^{g\,*} / G$ for every trader $g \in \{1, ..., G\}$. Let us assume that z^* is *such that there exists two traders* g *and* g' *with* $q_{jit}^{g\,*} \geq q_{jit}^{g'*} > 0$, *the conditions* (B.4) *for these two traders are such that:*

Trader g *:* $$P_{it}^{*} - P_{jt}^{*} - T_{jit} - \left(\frac{d_i\, b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j} \right) q_{jit}^{g\,*} - \xi_{jit}^{*} = 0\,, \tag{B.10}$$

Trader g'*:* $$P_{it}^{*} - P_{jt}^{*} - T_{jit} - \left(\frac{d_i\, b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j} \right) q_{jit}^{g'*} - \xi_{jit}^{*} = 0\,. \tag{B.11}$$

Subtracting (B.10) from (B.11), we obtain the equation $\left(\frac{d_i\, b_i}{b_i + d_i} + \frac{d_j b_j}{b_j + d_j} \right)\left(q_{jit}^{g'*} - q_{jit}^{g\,*} \right) = 0$ *that can only be verified if* $q_{jit}^{g\,*} = q_{jit}^{g'*}$. *As such a reasoning is valid for every pair of traders, we obtain* $q_{jit}^{g\,*} = Q_{jit}^{g\,*} \big/ G$ *for every trader* $g \in \{1,...,G\}$. *As the condition* $q_{jit}^{g\,*} = Q_{jit}^{g\,*} \big/ G$ *for every trader* $g \in \{1,...,G\}$ *also holds when the aggregate output is equal to 0, we can simply replace* $q_{jit}^{g\,*}$ *by* $Q_{jit}^{g\,*} \big/ G$ *in the equilibrium conditions* (B.4) *and* (B.5) *to prove that if these conditions hold so do the aggregate complementarity conditions* (B.6) *and* (B.7). *Q.E.D.*

Appendix C – A dynamic specification

The specification in Section 3.2 has a static nature. This technical Appendix outlines how it can be extended to model the dynamics of the inter-period linkages that may exist in commodity markets. Section C.1. recalls how the analysis in Kleit (2001) can be adapted to correct for serial correlation. Section C.2. explains how this approach can also be adapted to model a possibly time-varying variance for the residual in regimes I and II.

C.1 – Correcting for autocorrelation

Serial correlation due to both supply shocks and speculative storage activity is commonly observed in the empirical studies dedicated to commodity prices (Deaton and Laroque, 1996).[13] As the presence of unmodeled autocorrelation can result in inefficient estimates, the presence of serial correlation has to be appropriately corrected for.[14] Interestingly, Kleit (2001) details a relevant strategy to overcome this limitation and adjust for the possible presence of serial correlation in our error term ε_{jit}. One has to keep in mind that the exact value $\varepsilon_{ji(t-1)}$ cannot be directly observed. However, one can

[13] *In the application discussed in this paper, two arguments motivate the presence of autocorrelation. First, a pipeline system can be described as a slow-moving transportation infrastructure because a couple of hours are needed to move a given molecule of methane from one market to the other. Second, the operation of a natural gas pipeline system creates a temporary energy storage (the so-called line-pack buffer). As a result, daily observations are likely to jointly represent the outcome of decisions taken both today and yesterday.*

[14] *Barrett and Li (2002, footnote 3) discussed the serial correlation issue and claimed that the Cochrane-Orcutt method could be used to correct for serial correlation. However, the distribution of the observed residuals is dramatically modified from one observation to the next in case of a regime switch. Therefore, one may question the validity of a Cochrane-Orcutt approach.*

consider the expected value of $\varepsilon_{ji(t-1)}$, given the evidence provided by the previous observation, which results in the modified specification:

Regimes I & II: $$R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right) = \varepsilon_{jit} \quad \text{(C.1)}$$

Regimes III$_a$, III$_b$ & IV: $$R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right) = \varepsilon_{jit} + \mu_{jit} \quad \text{(C.2)}$$

Regimes V & VI: $$R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right) = \varepsilon_{jit} - \upsilon_{jit} \quad \text{(C.3)}$$

where: ρ_{ji} is an autocorrelation coefficient such that $-1 < \rho_{ji} < 1$; $\eta_{ji(t-1)}$ is the observed lagged residual, that is $\eta_{ji(t-1)} \equiv \pi_{ji(t-1)} - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-2)}\middle|\eta_{ji(t-2)}\right)$; and $\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right)$ represents the expected value of $\varepsilon_{ji(t-1)}$ given evidence provided by the observed lagged residual.

The expected value $\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right)$ can be computed as follows. Given the observed value of the lagged residual $\eta_{ji(t-1)}$ and the parameter vector $\theta_1 \equiv \left(\theta, \rho_{ji}\right)$, the probability $\mathrm{P}_{t-1}^r \equiv \mathrm{P}_{t-1}\left(r\middle|\eta_{ji(t-1)}, \theta_1\right)$ that the residual observed at time $t-1$ was generated by regime r is (Kiefer, 1980; Spiller and Wood, 1988, p.889–90):

$$\mathrm{P}_{t-1}^r = \frac{\lambda_r f_{ji(t-1)}^r\left(\eta_{ji(t-1)}\middle|\theta_1\right)}{\left(\lambda_{III_a} + \lambda_{III_b}\right) f_{ji(t-1)}^{III}\left(\eta_{ji(t-1)}\middle|\theta_1\right) + \sum_{\substack{k=I \\ k\neq III}}^{VI} \lambda_k f_{ji(t-1)}^k\left(\eta_{ji(t-1)}\middle|\theta_1\right)}. \quad \text{(C.4)}$$

The expected value $\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right)$ can be constructed from the observed residual $\eta_{ji(t-1)}$ by: (i) subtracting $\mathrm{E}(\mu)$ the expected value of the one-sided random variable μ_{jit} weighted by the probability to observe the regimes III$_a$, III$_b$ or IV; and, (ii) adding $\mathrm{E}(\upsilon)$ the expected value of the non-negative half-normal random variable υ_{jit} weighted by the probability to observe the regimes V or VI,[15] that is:

$$\mathrm{E}\left(\varepsilon_{ji(t-1)}\middle|\eta_{ji(t-1)}\right) = \eta_{ji(t-1)} - \left[\mathrm{P}_{t-1}^{III_a} + \mathrm{P}_{t-1}^{III_b} + \mathrm{P}_{t-1}^{IV}\right]\mathrm{E}(\mu) + \left[\mathrm{P}_{t-1}^{V} + \mathrm{P}_{t-1}^{VI}\right]\mathrm{E}(\upsilon). \quad \text{(C.5)}$$

[15] *Denoting* ϕ *the density function of the standard normal distribution and* Φ *its cumulative distribution function, these expected values are:* $\mathrm{E}(\mu) \equiv \sigma_\mu \phi(0)/(1-\Phi(0))$ *and* $\mathrm{E}(\upsilon) \equiv \sigma_\upsilon \phi(0)/(1-\Phi(0))$.

The construction of $\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right)$ can be nested within the likelihood specification above. So, the estimation proceeds again from a maximization of the log-likelihood function with respect to the regime probabilities λ and the parameters θ_1 subject to the preceding constraints and to $-1 < \rho_{ji} < 1$.[16]

C.2 – An adapted GARCH specification

Regimes I and II model the cases of zero marginal profit to spatial arbitrage. In these regimes, the random variable representing the marginal profit to spatial arbitrage is assumed to be equal to the stochastic error term ε_{jit} which has the same finite variance σ_ε^2 for all observations. Yet, one may question the relevance of this homoscedastic assumption as a large empirical literature has documented the tendency of commodity prices to exhibit time-varying volatilities. Accordingly, the spatial price differential (and thus the marginal rents to spatial arbitrage) is likely to show signs of heteroscedasticity. Inspired by the strategy proposed in Kleit (2001) to correct for serial correlation, it is also possible to design a modified specification whereby the variance of the marginal rents to spatial arbitrage observed in regimes I and II is allowed to vary over time.

For the purpose of capturing the dynamics of uncertainty, a Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model (Bollerslev, 1986) represents an attractive approach that has been widely applied to model commodity markets. Given the time series R_{jit} and Q_{jit} defined above, a GARCH(1,1) specification can be written as follows:

Regimes I & II: $$R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right) = \varepsilon_{jit} \tag{C.6}$$

Regimes $\mathrm{III_a}$, $\mathrm{III_b}$ & IV: $$R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right) = \varepsilon_{jit} + \mu_{jit} \tag{C.7}$$

Regimes V & VI: $$R_{jit} - \alpha_{ji} - X_t^{'}\beta_{ji} - Q_{jit}\gamma - \rho_{ji}\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right) = \varepsilon_{jit} - \upsilon_{jit} \tag{C.8}$$

$$\varepsilon_{jit} = h_{jit} e_{jit} \tag{C.9}$$

$$h_{jit}^2 = \varpi_{ji} + \delta_{ji}\left[\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right)\right]^2 + \varphi_{ji} h_{ji(t-1)}^2 \tag{C.10}$$

where: (C.6), (C.7) and (C.8) are the mean equations; (C.10) is the conditional variance equation; (C.9) relates the random error ε_{jit} to the standardized residual e_{jit} which is assumed to be an i.i.d. standard normal random variable; and ϖ_{ji}, δ_{ji} and φ_{ji} are the usual, non-negative, GARCH(1,1) parameters.

---

[16] *Regarding the particular case of the first observation, an arbitrary value has to be taken for* $\mathrm{E}\left(\varepsilon_{ji0} \middle| \eta_{ji0}\right)$ *because* η_{ji0} *cannot be observed. In this paper, the initial value* $\mathrm{E}\left(\varepsilon_{ji0} \middle| \eta_{ji0}\right)$ *is taken as equal to zero (that is, the conditional mean of* ε_{jit} *given* $\varepsilon_{ji(t-1)}$*).*

Compared to the usual GARCH specification, equation (C.10) involves the use of the squared expected value $\left[\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right)\right]^2$ in spite of the true value $\varepsilon^2_{ji(t-1)}$ which cannot be observed in this regime switching model. Again, the construction of the expected value $\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right)$ is based on the following reasoning: given the observed value $\eta_{ji(t-1)}$, the values of the parameters $\theta_2 \equiv \left(\theta_1, \varpi_{ji}, \delta_{ji}, \varphi_{ji}\right)$ and $h^2_{ji(t-2)}$, it is possible to evaluate the probabilities P^r_{t-1} and thus $\mathrm{E}\left(\varepsilon_{ji(t-1)} \middle| \eta_{ji(t-1)}\right)$ using (C.5).