Bocconi

WORKING PAPER SERIES
ISSN 2785-1079

Number 01
DECEMBER 2018

CONVERGENCE OF EUROPEAN NATURAL GAS PRICES

Andrea Bastianin
Marzio Galeotti
Michele Polo

Università Bocconi
GREEN
Centre for Geography, Resources, Environment, Energy and Networks

Convergence of European natural gas prices

Andrea Bastianin
University of Milan–Bicocca

Marzio Galeotti
University of Milan
and GREEN–Bocconi

Michele Polo
Bocconi University
and GREEN–Bocconi

December 6, 2018

Abstract: Over the period 2015–2050 the consumption of natural gas of European OECD countries is expected to grow more than the consumption of any other energy source. Although these countries are interconnected and in most cases share a common currency, their wholesale national gas markets are highly heterogeneous. We study the determinants of cross–country convergence of natural gas prices for industrial consumers in fourteen European countries. Our empirical analysis is based on the notions of pairwise, relative and σ–convergence. We show that there is evidence of pairwise price convergence and that some key characteristics of gas markets, such as the maturity of trading hubs and the degree of interconnection, are positively associated with the existence of a convergence process. This result carries over to the notion of σ–convergence and is robust to a number of changes in the implementation of the statistical tests. The analysis of relative convergence points to the existence of price–growth convergence, while price–level convergence is not supported by the data. Lastly, we assess the the short-run implications of price convergence focusing on the speed of reversion to equilibrium after a system–wide shocks hits the cointegrating relation.

Key Words: convergence; natural gas price; trading hub; σ–convergence; relative convergence.
JEL Codes: C22; L95; Q02; Q35; Q41.

Corresponding author: Andrea Bastianin, Department of Economics, Management, and Statistics, University of Milan – Bicocca, Via Bicocca degli Arcimboldi, 8, Building U7, 20126 - Milano – Italy. Email: `andrea.bastianin@unimib.it`.

1 Introduction

In 2015 natural gas accounted for 22% of the total primary energy consumption of OECD European countries and over the 2015–2050 period is expected to grow more than the consumption of any other energy source (EIA, 2017).[1] Given its strategic importance, the functioning of the natural gas market is high on the agenda of European regulators who have devoted considerable effort to the creation of a single market for energy at least since the Single European Act of 1986. Three consecutive legislative packages were subsequently adopted between 1996 and 2009 with the aim of harmonizing and liberalizing the EU's internal energy market. As a result of these measures, new gas and electricity suppliers can enter the Member States' markets, while both industrial and domestic consumers are now free to choose their own suppliers. The appearance of a multitude of market operators that need to balance their positions has also prompted the development of trading hubs in several European countries (Heather, 2012; Miriello and Polo, 2015; Hulshof et al., 2016; del Valle et al., 2017). Whether regulatory reforms and the development of trading hubs have contributed to the integration of national markets and to the alignment of gas prices for both industrial and residential use is a highly debated topic (see e.g. Asche et al., 2017; Brau et al., 2010; Cremer and Laffont, 2002, and references therein).

In this paper we study the cross–country convergence of natural gas prices for industrial consumers in fourteen European countries relying on time series econometric techniques. These countries belong to the European Union, have interconnected natural gas markets, and in most cases share a common currency; however, their degree of interconnection and their wholesale national gas markets and trading hubs – where these exist – are highly heterogeneous in terms of maturity. We assess the association between the convergence of natural prices for industrial consumers, countries' characteristics, trading hub maturity and other institutional features.

Several papers have analyzed the impact of liberalizations on residential and industrial prices focusing on the process of integration and convergence across different locations. The methodology typically relies on the assessment of the Law of One Price (LOP) using cointe-

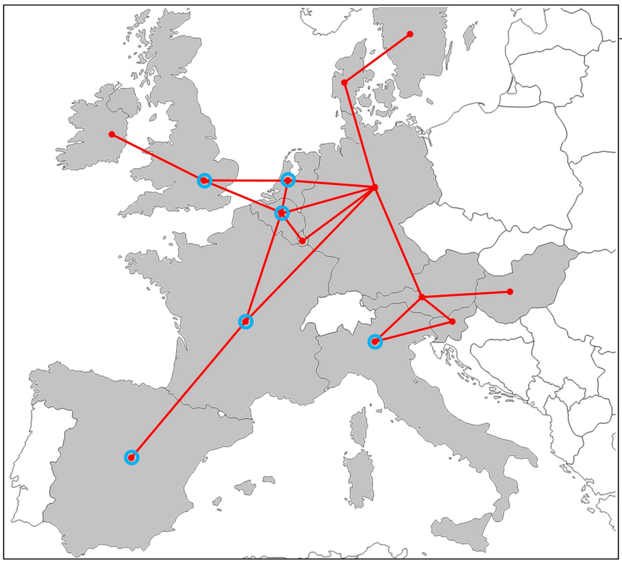
[1] In 2015 natural gas consumption of OECD Europe represented 13% of world consumption.

gration analysis. In case of full integration of two markets, industrial consumers should pay the same price, once transaction and transportation costs are accounted for. For historical reasons, this strand of the literature has focused first on North America (see e.g., De Vany and Walls, 1993; King and Cuc, 1996; Serletis and Herbert, 1999; Cuddington and Wang, 2006; Park et al., 2008; Apergis et al., 2015) where in the mid-80's governments implemented several policies aimed at deregulating the market for natural gas. More recent are the contributions focusing on European gas markets (see e.g., Asche et al., 2002; Neumann et al., 2006; Robinson, 2007; Renou-Maissant, 2012; Growitsch et al., 2015; del Valle et al., 2017) or presenting international comparisons (Siliverstovs et al., 2005; Li et al., 2014).

This study is closely related to two papers. Robinson (2007) focused on annual retail natural gas prices for six EU Member States and showed that over the 1978–2003 period there is evidence of β-convergence, as well as of convergence toward the group average using the test proposed by Bernard and Durlauf (1995, 1996). Renou-Maissant (2012) tested the LOP and analyzed convergence across six European natural gas markets. Relying on half–yearly data for industrial consumers over the 1991–2001 period the author showed that there is an on-going process of price convergence, but that the strength of market integration varies through time and across countries. Compared with these analyses, we consider a larger group of countries, use the less stringent concept of "pairwise convergence" (Pesaran, 2007) and rely on different econometric methods that have some advantages over those used by these authors. Since "pairwise convergence" is linked with the notion of σ-convergence, we present empirical evidence also on this issue. In addition, we consider the notion of relative convergence due to Phillips and Sul (2007) and we assess how convergence affects the speed of adjustment to equilibrium after a shock. Lastly, we study the association between the existence of price convergence and key characteristics of natural gas markets. Overall, we find evidence convergence for the price paid by industrial consumers, in line with Robinson (2007) and Renou-Maissant (2012). The strength of this result however depends on the characteristics of national gas markets, as well as on the existence and the maturity of gas hubs and on the degree of interconnection of the markets.

The rest of the paper is organized as follows. In Section 2 we describe the data and provide some background on the European gas market. Section 3 introduces the economet-

Figure 1: List of countries, hubs, LNG facilities and interconnections as of 2017



ISO	Country	Hub
AT	Austria	CEGH
BE	Belgium	ZTP
DE	Germany	VHP NCG, VHP-GASPOOL
DK	Denmark	GTF, NPTF
ES	Spain	VIP PIRINEOS, PVB
FR	France	PEG NORD, TRS
HU	Hungary	MPG
IE	Ireland	IBP
IT	Italy	PSV
LU	Luxembourg	-
NL	Netherlands	TTF
SE	Sweden	-
SI	Slovenia	-
UK	United Kingdom	NBP

Notes: lines in the map sketch the interconnections between natural gas markets. Blue circles signal the presence of one or more Liquefied Natural Gas regasification facilities. Authors' elaborations using the 2017 ENTSOG Capacity map dataset (`https://www.entsog.eu`).

ric methodology. Section 4 presents the main empirical results. Section 5 discusses some extensions and robustness checks. Section 6 concludes. Further results and methodological details are provided in the Appendix.

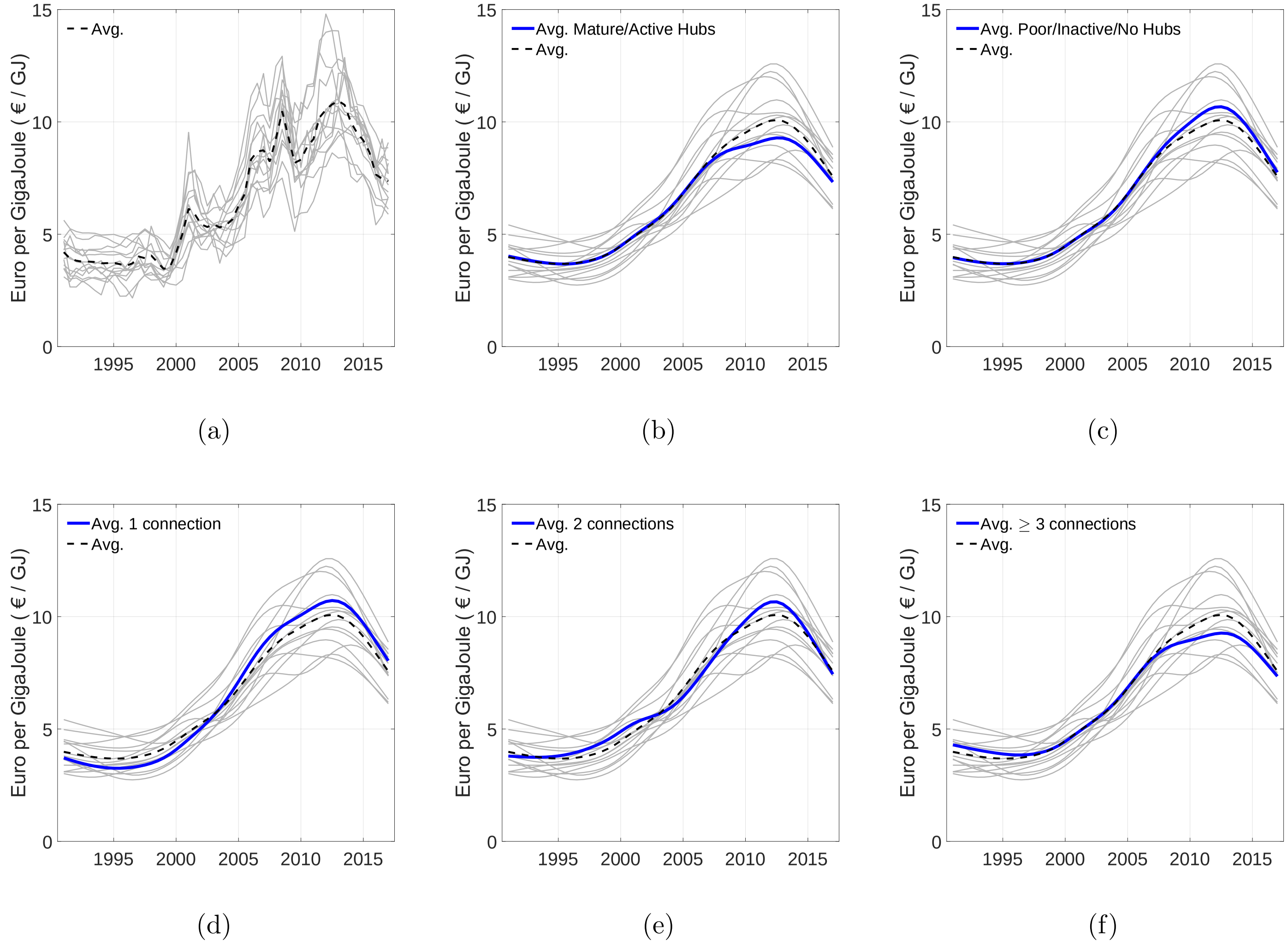
2 EU gas markets: data and background aspects

2.1 Data

We consider natural gas prices paid by industrial consumers in fourteen European countries, namely: Austria, Belgium, Germany, Denmark, Spain, France, Hungary, Ireland, Italy, Luxembourg, Netherlands, Sweden, Slovenia, United Kingdom. In line with previous studies we rely on before tax prices to avoid that fiscal policy might act as confounding factor in our convergence analysis (see e.g., Robinson, 2007; Renou-Maissant, 2012). As shown in Figure 1, each of these markets is interconnected with at least one of the other countries, which is the reason why the above countries were selected (see also Figure A1 in Appendix A.1).

We have sourced half–yearly natural gas prices for industrial consumers belonging to the medium consumption band (i.e., entities with consumption of 10,000 – 100,000 gigajoule per year) from Eurostat. The sample period runs from the first semester of 1991 through

Figure 2: Natural gas prices, hub maturity & number of connections: $1991h1 - 2017h1$



Notes: Panel (a) displays price series, while in panel (b-f) all lines are price trends estimated with the Hodrick and Prescott (1997) filter with smoothing parameter equal to 100, as suggested by the frequency power rule of Ravn and Uhlig (2002). Thin gray lines represent natural gas prices in each of the 14 European countries; dark dashed lines are the sample average of prices for the 14 countries. In panel (b-c) thick dark continuous lines indicate the average price in "Mature/Active Hubs" ("Poor/Inactive No Hubs") in the classification of countries based on hub maturity (see EFET, 2016; Heather and Petrovich, 2017). In panel (d-f) thick dark continuous lines identify the average price in hubs with a given number of connections (i.e. 1, 2, ≥ 3).

the first semester of 2017. Denoting semesters as "h", our data span the 1991:h1 – 2017:h1 period, for a total of 53 observations per country.[2]

Figure 2(a) shows that, that consistently with the notion that European gas markets are integrated, prices for industrial consumers are highly correlated; however, it is impossible to spot clear signs of convergence. Moreover, the spread of the series around the sample average tends to increase during episodes of high crude oil price volatility. This happens for instance in 2001, 2007/08 and 2012 and supports empirical studies showing that the gas pricing mechanism is still, at least to some extent, influenced by what happens in the crude oil market (Asche et al., 2013; Nick and Thoenes, 2014). In Figures 2(b-f) we show that, at least since 2007, industrial consumers in countries with more than tree interconnections or that host the most active trading hubs have experienced prices lower than the EU average.

2.2 European markets for natural gas: background

The analysis of price convergence for industrial gas consumers requires to understand how prices in European markets are set, how the inter–relation among them takes place and affects the price setting mechanisms. While in continental Europe gas trading is still largely based on long–term contracts indexed to the price of crude oil, there is evidence that gas–specific factors are becoming increasingly important (see e.g., Siliverstovs et al., 2005; Asche et al., 2017). Figure 1 shows that as of 2017 all countries in our sample — except Luxembourg, Slovenia and Sweden — hosted at least one gas trading hub. Miriello and Polo (2015) provide a theoretical framework to analyze the patterns of development of wholesale gas markets and their relationship with the liberalization processes.[3]

[2]Data are available from 1985, but before 1991 the effective sample size varies greatly and is often significantly shorter. Since the time series for Austria, Hungary, Sweden, and Slovenia start between 1991:h1 and 1996:h1, we rely on the average growth rate of prices for the remaining countries to backcast missing observations. Let $p_{1,t} = \log(P_{1,t})$ be the log–price for Sweden for t =1996:h1, ..., 2017:h1 and let $\overline{\Delta p_t} = (1/10)\sum_{j=1}^{10} \log(P_{j,t}/P_{j,t-1})$ be the average price growth rate for Belgium, France, Germany, Denmark, Spain, Ireland, Italy, Luxembourg, Netherlands and United Kingdom. Then $P_{1,t-1} = \exp^{\overline{\Delta p_t}} P_{1,t}$ for t =1996:h1, ..., 1991:h2.

[3]According to Miriello and Polo (2015) the liberalization has promoted a progressive fragmentation of the different market segments; consequently each operator manages smaller volumes and narrower portfolios of contracts. Then, the ability to compensate the gap between demand and supply of each individual contract by compensating imbalances of different sign within the portfolio is reduced. Balancing the system through direct trade among operators with different net positions has become a priority. The wholesale market in its initial phase has therefore developed to cope with these balancing needs. A more fragmented and more

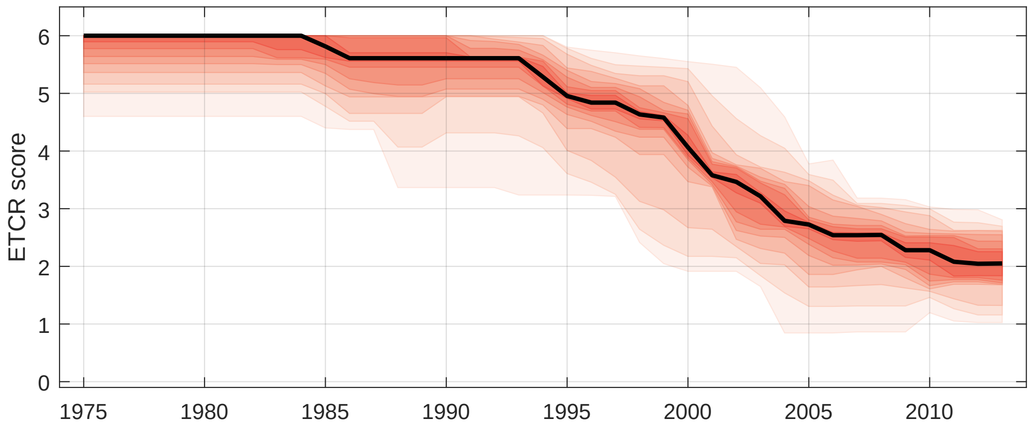
The process of liberalization of EU natural gas markets — led by the UK experience (Asche et al., 2013) — has involved three main legislative packages. With the First Gas Directive, promulgated in 1998 (1998/30/EC), gas markets were opened up to competition by facilitating the entry in the competitive segments of the industry. New common rules for transmission, distribution, supply and storage of natural gas were adopted. The Second Gas Directive (2003/55/EC) provided for the unbundling of the vertically integrated gas operators and made the transport networks of gas independent from production and supply. Industrial consumers were allowed to choose their suppliers since July 2004, while for household consumers the date was delayed to July 2007. The Third Gas Directive (2009/73/EC) improved the functioning of the internal energy market and resolved structural problems, plus unbundling of energy suppliers from network operators.[4]

The dynamics of the liberalization process over the 1991–2017 period is illustrated in Figure 3, where we plot the OECD's Energy, Transport and Communications Regulation (ETCR) index for the natural gas markets of the countries in our sample (for details see Koske et al., 2015, and Appendix A.2.). The ETCR index takes on values between zero and six, with lower values indicating fewer restrictions to competition. As we can see, a downward sloping trend is clearly visible in Figure 3; moreover, column 7 of Table 1 shows that in 2013 the ETCR index for natural gas did not exceed three for any market and was equal to zero for the UK, the country that started the liberalization process first. Column 6 shows that also the Product Marked Regulation (PMR) index — an overall score for natural gas, electricity, air, rail, road transport, post and telecommunications — is always below three, indicating that reforms aimed at liberalizing network industries have been enacted by the governments of the countries in our sample.

liquid wholesale market, in turn, has reduced the ability of large operators to manipulate the price, that has become a more reliable signal. These processes have made the wholesale markets an appealing second source of gas purchase as an alternative to long-term contracts, further pushing up liquidity. Lastly, price variability has required to manage price risk by developing financial instruments, the third stage of wholesale gas markets. See Miriello and Polo (2015) for details and evidence.

[4]More recent EU legislation affecting natural gas markets includes: the Proposal for a Regulation "concerning measures to safeguard the security of gas supply and repealing Regulation (EU) No 994/2010" (COM(2016)52/F1) and the "Clean Energy For All Europeans" package, known as the "Winter Package", published on 30 November 2016, consisting of numerous legislative proposals together with accompanying documents, aimed at further completing the internal market for electricity and at implementing the Energy Union.

Figure 3: OECD's Energy, Transport and Communications Regulation (ETCR) index for natural gas: 1991-2017.



Notes: the ETCR index for natural gas shown in the top panel is the simple average of the four sub-indexes covering different dimensions of market reforms: entry regulation, public ownership, vertical integration and market structure. Data represent a scoring system on a scale from zero to six, where values near zero indicate fewer restrictions to competition. Since the latest release of the ETCR database ends in 2013, for each country we assign the 2013 value to the 2014-17 period. The dark line represents the median value of the index, the shaded area is delimited by the 10th and 90th percentile of the distribution of the index across countries. Source: Authors' elaboration on data from Koske et al. (2015).

Figure 3 also highlights that the distribution of ETCR index for natural gas is highly dispersed around the median value over the whole period, thus suggesting that some heterogeneity in the level of competitiveness of different gas markets tends to persist. This fact is mirrored in the heterogeneity that characterizes the degree of maturity of European gas hubs. Following Heather and Petrovich (2017) these can be described according to "Score" and "Category" in Table 1, columns 3 and 4. The subjective scoring system in column 3 evaluates three elements: commercial acceptance, political willingness and cultural attitudes to trading.[5] Higher scores identify mature hubs, while lower scores are associated to less active hubs. It is seen that mature hubs are hosted in the UK, Netherlands, France and Germany.

To sum up, more mature and liquid wholesale markets allow to manage efficiently balancing needs, adjusting internal imbalances of operators and the overall imbalances of the system, reacting quickly to shocks, preventing price manipulations and providing reliable price signals of the state of the system. In these markets arbitrage opportunities within the

[5]For details see Heather and Petrovich (2017). Interestingly, this classification is consistent with that provided by EFET (2016) based on market statistics related with depth, liquidity and transparency of gas hubs (see Heather and Petrovich, 2017, Table 5).

Table 1: Descriptive statistics for the European natural gas markets in 2017

		Hub			ETCR score	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
iso	country	Score	Category	Connections	PMR	Gas
AT	Austria	13.5	Poor	4	1.5	2.2
BE	Belgium	17.5	Active	5	1.8	1.7
DE	Germany	19	Mature	6	1.3	1.2
DK	Denmark	14	Poor	2	1.6	2.6
ES	Spain	13.5	Poor	1	1.6	1.1
FR	France	18.5	Mature	3	2.5	2.5
HU	Hungary	9	Poor	1	1.7	1.7
IE	Ireland	—	—	1	2.2	3.0
IT	Italy	15	Active	2	2.0	1.9
LU	Luxembourg	—	—	2	2.7	2.6
NL	Netherlands	19.5	Mature	3	1.6	2.3
SE	Sweden	—	—	1	1.9	1.7
SI	Slovenia	—	—	2	2.9	2.8
UK	UK	20	Mature	3	0.8	0.0

Notes: columns 3, 4, 5 indicate the 2016 classification of countries based on hub maturity (see EFET, 2016; Heather and Petrovich, 2017), where "—" denotes that there are no information on the relevant hub or that there is no hub. Columns 6 and 7 report indices sourced from the OECD's Energy, Transport and Communications Regulation (ETCR) dataset (Koske et al., 2015). These indices represent a scoring system on a scale from zero to six, where values near zero indicate fewer restrictions to competition. Column 6 shows the Product Market Regulation index ("PMR"); this is an overall score for seven network industries: gas, electricity, air, rail, road transport, post and telecommunications. Column 7 shows the ETCR index for the natural gas market ("Gas").

markets are exploited by operators leading to a rapid convergence to the market equilibrium. If the different national gas systems were completely isolated with no interconnection, but very liquid, we should observe prices to reflect country–specific fundamentals and idiosyncratic shocks. So long as these factors are different across countries, prices should not necessarily converge. Therefore, the relationship among national gas systems might also contribute to price convergence. These are affected by the level of interconnection that takes place through the international pipelines and the LNG terminals that allow intra-community trades among member countries.

Casual inspection of columns 3 and 5 of Table 1 suggests that hub maturity is positively correlated with the number of interconnections of a given gas market with the other countries in the sample. The map in Figure 1 shows that only 6 out 14 countries had an operating Liquefied Natural Gas regassification facility in 2017. Appendix A.1 provides additional information on the technical physical capacity of European natural gas markets.

3 Pairwise price convergence

3.1 Theoretical background

We analyze cross–country price convergence relying on log–price differentials:

$$d_{ij,t} = \log\left(\frac{P_{i,t}}{P_{j,t}}\right) = p_{i,t} - p_{j,t} \qquad \text{for } i,j = 1,\ldots,N \text{ and } i \neq j \tag{1}$$

where $p_{i,t} \equiv \log P_{i,t}$ and $P_{i,t}$ is the before tax price of natural gas for industrial consumers in country i. Bernard and Durlauf (1995, 1996) proposed the following definition of convergence:

$$\lim_{H\to\infty} E\left(p_{i,t+H} - p_{j,t+H} \,|\, \mathcal{I}_t\right) = 0 \qquad \text{for } H > 0 \tag{2}$$

where $\mathcal{I}_t$ is the information set at time t containing current and past information on prices. In this setting, price convergence requires that the long–term forecast of $d_{ij,t}$ tends to zero as the forecast horizon H increases. This implies that a necessary, but not sufficient condition for prices in countries i and j to converge, is that they are cointegrated with cointegrating vector $[1,\ -1]'$. Suppose that the log–price of natural gas in each country can be written as:

$$p_{i,t} = \gamma_i + \beta_i t + \psi_{i,t} \qquad \text{for } i = 1,\ldots,N \tag{3}$$

Equation (3) expresses $p_{i,t}$ as the sum of a country fixed effect (γ_i), a deterministic trend component ($\beta_i t$) and an error term ($\psi_{i,t}$) that can be either integrated of order one, $I(1)$, or stationary. If prices in country i and j are cointegrated with cointegrating vector $[1,\ -1]'$, there exists a linear combination $z_t = p_{i,t} - p_{j,t}$ that is stationary or trend stationary. The convergence condition in Equation (2) can therefore be written as the limit of:

$$E\left(p_{i,t+H} - p_{j,t+H} \,|\, \mathcal{I}_t\right) = (\gamma_i - \gamma_j) + (\beta_i - \beta_j)(t+H) + E\left(\psi_{i,t+H} - \psi_{j,t+H} \,|\, \mathcal{I}_t\right) \tag{4}$$

If $\psi_{i,t}$ and $\psi_{j,t}$ are zero–mean independent stationary processes, it follows that $\lim_{H\to\infty} E\left(\psi_{i,t+H} - \psi_{j,t+H} \,|\, \mathcal{I}_t\right) = E\left(\psi_{i,t} - \psi_{j,t}\right) = 0$. In this case, natural gas prices in country i and j converge if $\gamma_i = \gamma_j$ and $\beta_i = \beta_j$. If instead $\psi_{i,t}$ and $\psi_{j,t}$ are $I(1)$, we also require the two prices to share a common stochastic trend, that is: $\psi_{i,t} = \psi_{j,t}$. These conditions imply that economies i and j are equal almost in every respect (i.e., the "poolability" restriction: $\gamma_i = \gamma_j$), that the two price series share a deterministic trend (i.e., the "cotrending" restriction: $\beta_i = \beta_j$) and, in case of a unit root in the price series, that they are cointegrated ($\psi_{i,t} = \psi_{j,t}$). Of these conditions the "poolability" restriction ($\gamma_i = \gamma_j$), is the most unlikely to be satisfied. In addition, assessing cointegration involves a pre–test bias due to the fact that the individual price series need to be preliminary tested for the presence of a unit root. However, if prices are generated by a "near unit root process", standard unit–root tests have low power against the alternative hypothesis and hence lead to biased second–stage inferences (Cavanagh et al., 1995; Elliott, 1998).

Pesaran (2007) showed that a less stringent formulation of convergence is based on the conditional probability of observing an arbitrarily small log–price differential. The concept of "pairwise convergence" implies that prices in country i and j converge if:

$$\Pr\left(|p_{i,t+H} - p_{j,t+H}| < C \mid \mathcal{I}_t\right) > \pi \qquad \text{for } C > 0,\ 0 \leq \pi < 1,\ \forall H > 0 \tag{5}$$

Condition (5) is satisfied if the log–price differential does not display stochastic nor deterministic trends, that is $\beta_i = \beta_j$ and $\psi_{i,t} = \psi_{j,t}$ if prices are $I(1)$. However, "pairwise convergence" does not require "poolability" ($\gamma_i = \gamma_j$), does not rely on unit root tests for the individual price series and hence allows to eschew pre–test issues. In fact, this notion of convergence allows two countries to be different (with country heterogeneity captured by γ_i and γ_j) and requires only testing for the absence of unit roots and linear deterministic trends in the log–price differential. Extension to a multi–country setting requires pairwise convergence across all country pairs.[6]

[6]In a multi–country setting condition (5) is (see Pesaran, 2007):

$$\Pr\left\{\bigcap_{i=1,\ldots,N-1\ j=i+1,\ldots,N} |p_{i,t+H} - p_{j,t+H}| < C \middle| \mathcal{I}_t\right\} > \pi \text{ for } C > 0,\ 0 \leq \pi < 1,\ \forall H > 0.$$

3.2 Tests of pairwise price convergence

Tests for pairwise convergence involve two distinct steps. First, since two prices converge if $d_{ij,t}$ is stationary with a constant mean, we need to test for the presence of a unit root in the log–price differential across all country pairs. Next, in the case of rejection of the null hypothesis of a unit root in $d_{ij,t}$, we check the cotrending condition, namely $\beta_i = \beta_j$. This is carried out, with an OLS regression of $d_{ij,t}$ on a constant and a linear trend. If the trend is not statistically distinguishable from zero, we conclude that prices in market i and j converge. Our baseline results are based on three tests for a unit root in $d_{ij,t}$, namely the Augmented Dickey–Fuller test (Dickey and Fuller, 1979) (ADF), the Generalized Least Squares Dickey–Fuller test (Elliott et al., 1996) (DF–GLS) and the Augmented Dickey–Fuller Weighted Symmetric test (Park and Fuller, 1995) (ADF–WS). The DF–GLS and the ADF–WS tests have been shown to be superior to the ADF test (Leybourne et al., 2005).

Denote by UR the test of the null hypothesis of a unit root $H_0 : d_{ij,t} \sim I(1)$ against the alternative that $d_{ij,t}$ is stationary, where UR is one of the three tests discussed above, that is $UR =$ ADF, DF–GLS, ADF–WS. Denoting the test carried out on observations $t = 1, \ldots, T$ as $UR_{ij,T}$ and its critical value of size α as $CV_{T,\alpha}$, the null hypothesis of "price divergence" is rejected if $UR_{ij,T} < CV_{T,\alpha}$. Noting that $\lim_{T\to\infty} \Pr\left(UR_{ij,T} < CV_{T,\alpha} \mid H_0\right) = \alpha$ and letting $Z_{ij,t} = 1$ if $UR_{ij,T} < CV_{T,\alpha}$, the fraction of the $N(N-1)/2$ pairs for which the unit root is rejected can be written as:

$$\overline{Z}_{NT} = \frac{2}{N(N-1)} \sum_{i=0}^{N-1} \sum_{j=1+i}^{N} Z_{ij,T} \tag{6}$$

Pesaran (2007) showed that $\overline{Z}_{NT}$ is a consistent estimator of α for large N and T, that is: $\lim_{T\to\infty} E\left(\overline{Z}_{NT} \mid H_0\right) = \alpha$. If price convergence is supported by the data, the null hypothesis of a unit root in the price differential should be rejected for a large number of country pairs: hence $\overline{Z}_{NT}$ should tend to unity and be much larger than the size of the test α. On the contrary, if the null of "price divergence" cannot be rejected for a large number of price differentials, $\overline{Z}_{NT}$ is expected to be close to the size of the test α.

4 Results

4.1 Pairwise convergence tests

Formal statistical tests of pairwise convergence are presented in Table 2. For the deterministic components of the unit root tests, we consider two cases: we include only the intercept ("const") or add also a linear trend, but only if it is significant at the 5% level ("const/trend"). We estimate the optimal number of lags included in the test equation with either Akaike (AIC) or Schwarz (SIC) Information Criterion. With $N = 14$ countries, we have a total of $N(N-1)/2 = 91$ log–price differentials to test. Log–price differentials are shown in Appendix A.3.

Pairwise price convergence is supported by the data when the null hypothesis is rejected a large number of times relative to the nominal size of the unit root test that in Table 2 is 5% or 10%. Columns 1, 3, 5 and 7 show that independently of the significance level, lag order, or exogenous variables included in the test equation, the fraction of rejections, $\bar{Z}_{NT}$, is always well above the nominal size of the test. At 5% (10%) significance level $\bar{Z}_{NT}$ ranges from 33.3% (37.9%) to 60.4% (72.5%). A high percentage of rejections of the null hypothesis of a unit root is however not enough to safely conclude that European gas prices have converged; in addition, log–price differentials should not feature any deterministic trend, but move around a constant mean. Columns 2, 4, 6 and 8 show the percent of log–price differentials for which the null hypothesis of a non–significant trend cannot be rejected. Student tests of the significance of the linear trend are conducted at the 5% significance level in a regression of $d_{ij,t}$ on a constant and a linear trend. Notice that the test is carried out only when the null hypothesis of unit root is rejected. There is evidence of convergence if the percentage of trend stationary series is relatively low (i.e. if the percentage in the "%t" column is high). It can be seen that in all cases such fraction never exceeds 50%, implying that the percentage of trend stationary series is relatively high. The existence of many trend stationary log–price differentials does not support convergence of prices across the countries in our sample.

Table 2: Pairwise convergence based on unit root tests: 1991:h1–2017:h1

	5%				10%			
	(1) AIC	(2) $\%t$	(3) SIC	(4) $\%t$	(5) AIC	(6) $\%t$	(7) SIC	(8) $\%t$
(a) ADF test								
const	37.4	44.1	35.2	50.0	56.0	39.2	45.1	46.3
const/trend	49.5	33.3	45.1	39.0	68.1	32.3	63.7	32.8
(b) DF–GLS test								
const	52.7	43.8	47.3	44.2	72.5	37.9	69.2	38.1
const/trend	52.7	43.8	48.4	43.2	70.3	39.1	64.8	40.7
(c) ADF–WS test								
const	60.4	40.0	60.4	40.0	72.5	37.9	72.5	37.9
const/trend	53.8	44.9	53.8	44.9	70.3	39.1	70.3	39.1

Notes: each panel shows in columns (1, 3, 5, 7) the percentage of the 91 log–price differentials ($d_{ij,t}$) for which the null hypothesis of unit root is rejected ($\bar{Z}_{NT}$). In the case of rejection of the null hypothesis (i.e. for stationary log–price differentials), columns (2, 4, 6, 8) show the percent of log–price differential for which the hypothesis of a non–significant trend is not rejected. Student tests of the significance of the linear trend are conducted at the 5% significance level in a regression of $d_{ij,t}$ on a constant and a linear trend (the test is carried out only when the null hypothesis of unit root is rejected). Convergence between prices is supported by the data when $\bar{Z}_{NT}$ is large relative to the significance level of the unit root test — 5% or 10% in this table — and the number of trend stationary series relatively low (i.e. the % in the "$\%t$" column is high). Each panel presents two cases for the deterministic component of the unit root test: "const" indicates that we included only the intercept; "const/trend" indicates that we included a linear trend only if it is significant at the 5% level. The lag length of test equations has been selected either with the Akaike or with the Schwarz information criterion, denoted as AIC and SIC, respectively. The maximum lag order is set equal to 4 that corresponds to two years. The three tests are the standard Augmented Dickey Fuller (ADF, Dickey and Fuller 1979), the DF–GLS of Elliott et al. (1996) and the ADF–WS of Park and Fuller (1995). Critical values for the ADF test are provided by MacKinnon (1996), while those for the DF–GLS and ADF–WS have been calculated by the authors using the response surface regressions in Cheung and Lai (1998) and Cheung and Lai (2009), respectively.

4.2 Pairwise convergence tests over different sample periods

Overall, Table 2 only partially supports the notion that industrial natural gas prices have converged: price differentials in most cases do not have a unit root, but many of them are stationary around a linear trend. We now analyze how pairwise convergence has evolved over time and in response to three significant events: the introduction of the Euro in 2002, the Second Gas Directive in 2004, the outbreak of the Great Recession and a change in Eurostat's data collection procedures in 2007.

Euro introduction — 2002. For several countries in our sample Euro coins and banknotes replaced national currencies and entered in circulation on 1 January 2002. Table 3 shows that both before and after 2002 the fraction of rejections of the unit root null hypothesis is much higher that the nominal size of the test. This result is robust across specifications and unit root tests. Interestingly, we can also observe that after the Euro was introduced the fraction of rejections always increases. Moreover, looking at columns 2, 4, 6 and 8 we see that the percentage of log–price differentials for which the null hypothesis of a non–significant

Table 3: Pairwise convergence tests – before & after Euro introduction

	5%				10%			
	(1) AIC	(2) %t	(3) SIC	(4) %t	(5) AIC	(6) %t	(7) SIC	(8) %t
(a) ADF test: 1991:h1–2001:h2								
const	20.9	78.9	25.3	73.9	27.5	80.0	35.2	71.9
const/trend	25.3	65.2	35.2	53.1	36.3	60.6	44.0	57.5
(b) ADF test: 2002:h1–2017:h1								
const	26.4	66.7	24.2	72.7	41.8	57.9	38.5	60.0
const/trend	33.0	53.3	30.8	57.1	50.5	47.8	48.4	47.7
(c) DF–GLS test: 1991:h1–2001:h2								
const	24.2	86.4	28.6	73.1	36.3	72.7	44.0	62.5
const/trend	31.9	65.5	35.2	59.4	40.7	64.9	51.6	53.2
(d) DF–GLS test: 2002:h1–2017:h1								
const	40.7	56.8	34.1	61.3	58.2	50.9	48.4	54.5
const/trend	45.1	51.2	42.9	48.7	60.4	49.1	57.1	46.2
(e) ADF–WS test: 1991:h1–2001:h2								
const	34.1	74.2	34.1	74.2	28.6	84.6	28.6	84.6
const/trend	27.5	92.0	27.5	92.0	30.8	78.6	30.8	78.6
(f) ADF–WS test: 2002:h1–2017:h1								
const	50.5	56.5	50.5	56.5	54.9	52.0	54.9	52.0
const/trend	47.3	60.5	47.3	60.5	59.3	48.1	59.3	48.1

Notes: The maximum lag order is set equal to 2 that corresponds to one year; for further details see notes to Table 2.

trend cannot be rejected is very high. Overall, Table 3 suggests that there is evidence of pairwise convergence especially after 2001.

The Second Gas Directive — 2004. With the Second Gas Directive of 2003 industrial consumers were allowed to freely choose their suppliers. Since the Directive entered into force in July 2004, we split the sample in two sub–periods: 1991:h1–2004:h1 and 2004:h2–2017:h1. Table 4 shows that after 2004 the percentage of rejections of the null hypothesis exhibits a uniform increase and so does the share of log–price differentials that do not display statistically significant linear trends. These results not only confirm that there is an overall convergence pattern that characterizes the countries in our sample, but also point to the existence of some association between the liberalization of the natural gas market and the degree of convergence. Note however that we cannot draw conclusions on the causality between liberalizations and convergence, in that there might be confounding factors we are not controlling for.

Table 4: Pairwise convergence tests – before & after the Second Gas Directive

	5%				10%			
	(1) AIC	(2) %t	(3) SIC	(4) %t	(5) AIC	(6) %t	(7) SIC	(8) %t
(a) ADF test: 1991:h1–2004:h1								
const	24.2	90.9	20.9	84.2	33.0	70.0	34.1	77.4
const/trend	29.7	74.1	23.1	76.2	34.1	67.7	36.3	72.7
(b) ADF test: 2004:h2–2017:h1								
const	25.3	69.6	23.1	61.9	36.3	54.5	31.9	55.2
const/trend	34.1	51.6	30.8	46.4	39.6	50.0	38.5	45.7
(c) DF–GLS test: 1991:h1–2004:h1								
const	23.1	71.4	31.9	79.3	34.1	61.3	44.0	72.5
const/trend	23.1	71.4	29.7	85.2	35.2	59.4	41.8	76.3
(d) DF–GLS test: 2004:h2–2017:h1								
const	38.5	54.3	30.8	53.6	50.5	56.5	42.9	56.4
const/trend	39.6	52.8	36.3	45.5	53.8	53.1	50.5	47.8
(e) ADF–WS test: 1991:h1–2004:h1								
const	31.9	65.5	31.9	65.5	31.9	65.5	31.9	65.5
const/trend	24.2	86.4	24.2	86.4	30.8	67.9	30.8	67.9
(f) ADF–WS test: 2004:h2–2017:h1								
const	50.5	56.5	50.5	56.5	50.5	56.5	50.5	56.5
const/trend	46.2	61.9	46.2	61.9	53.8	53.1	53.8	53.1

Notes: The maximum lag order is set equal to 2 that corresponds to one year; for further details see notes to Table 2.

Great Recession, Eurostat's methodology and more — 2007. There are three main reasons why in 2007 there might be a break in the natural gas price series. First, the National Bureau of Economic Analysis dates the start of the Great Recession in December 2007 and it is well known that the crude oil price rally in 2008/09 is one of the factors that has contributed to this event (Stock and Watson, 2012). Second, the rapid development of shale gas and shale oil production have affected the energy markets worldwide (Auping et al., 2016; Kilian, 2017; Koster and van Ommeren, 2015; Saussay, 2018). Lastly, a more practical concern is related with data collection procedures. Eurostat introduced a new methodology to collect natural gas price data in 2007. The new methodology uses "consumption bands" instead of "consumers standards". For the 1991:h1 – 2006:h2 period we use prices for I3–1 industrial consumers (i.e. annual consumption 41,860 gigajoule), while from 2007:h1 we use prices for Band I3 consumers (i.e., consumption of 10,000 – 100,000 gigajoule per year).

Table 5 shows evidence of pairwise convergence before and after 2007: although the fraction of rejections of the null hypothesis of a unit root in the log–price differential is quite

Table 5: Pairwise convergence based on unit root tests – before & after 2007

	5%				10%			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	AIC	$\%t$	SIC	$\%t$	AIC	$\%t$	SIC	$\%t$
(a) ADF test: 1991:h1–2006:h2								
const	15.4	64.3	23.1	66.7	25.3	60.9	30.8	64.3
const/trend	22.0	45.0	23.1	66.7	38.5	40.0	37.4	52.9
(b) ADF test: 2007:h1–2017:h1								
const	30.8	60.7	29.7	59.3	48.4	68.2	47.3	67.4
const/trend	37.4	50.0	35.2	50.0	54.9	60.0	54.9	58.0
(c) DF–GLS test: 1991:h1–2006:h2								
const	24.2	63.6	29.7	63.0	44.0	52.5	47.3	53.5
const/trend	26.4	58.3	30.8	60.7	42.9	53.8	42.9	59.0
(d) DF–GLS test: 2007:h1–2017:h1								
const	36.3	69.7	35.2	71.9	51.6	63.8	45.1	65.9
const/trend	41.8	60.5	41.8	60.5	53.8	61.2	49.5	60.0
(e) ADF–WS test: 1991:h1–2006:h2								
const	34.1	58.1	34.1	58.1	44.0	55.0	44.0	55.0
const/trend	28.6	69.2	28.6	69.2	40.7	59.5	40.7	59.5
(f) ADF–WS test: 2007:h1–2017:h1								
const	48.4	65.9	48.4	65.9	44.0	67.5	44.0	67.5
const/trend	44.0	72.5	44.0	72.5	45.1	65.9	45.1	65.9

Notes: The maximum lag order is set equal to 2 that corresponds to one year; for further details see notes to Table 2.

similar in the two sample periods, the percentage of rejections increases after 2007. This fact is consistent with the view that different factors and policy initiatives have contributed to a higher integration of European natural gas markets.

4.3 Pairwise convergence tests for different country groups

We now turn to the analysis of pairwise convergence across different country groups. We aggregate countries relying of three criteria: the existence of a trading hub, its maturity and the transmission capacity in 2017. The list of countries belonging to each group and its definition is provided in Appendix A.4.

Table 6 shows that the existence of a trading hub is associated with a small increase in the fraction of rejections of the null hypothesis of a unit root in the log–price differentials. Also the degree of development of the trading hub is associated with an increase in the share of country pairs for which there is evidence of pairwise convergence. Lastly, we group countries using their transmission capacity in 2017. The rows headed "Low transmission capacity"

Table 6: Pairwise convergence for different country groups: 1991:$h1$–2017:$h1$

(1)	(2) $\overline{Z}_{NT}$	(3) $\%t$
All	64.8	40.7
No Hub	63.9	52.2
Hub	65.5	33.3
Poor, Inactive or no hubs	63.2	41.7
Mature or Active hubs	73.3	36.4
Low transmission capacity	64.3	44.4
High transmission capacity	66.7	28.6

Notes: Column (1) shows which countries have been used in the computation of the pairwise convergence test, where: "All" means all countries, "No Hub" ("Hub") denotes countries without (with) a trading hub, "Poor, Inactive or No hub" ("Mature or Active hubs") indicates the classification countries based on hub maturity discussed in Section 2.1 (see EFET, 2016; Heather and Petrovich, 2017), "Low transmission capacity" ("High transmission capacity") indicates countries that have transmission capacity below (above) the 2017 median transmission capacity (source: 2017 ENTSOG Capacity map dataset:`https://www.entsog.eu`). Column (2) shows the percentage of log–price differentials ($d_{ij,t}$) for which the null hypothesis of unit root is rejected ($\bar{Z}_{NT}$) based on the DF–GLS test of Elliott et al. (1996) at the 10%. In the case of rejection of the null hypothesis (i.e. for stationary log–price differentials), column (3) shows the percent of log–price differential for which the hypothesis of a non–significant trend is not rejected. Student tests of the significance of the linear trend are conducted at the 5% significance level in a regression of $d_{ij,t}$ on a constant and a linear trend (the test is carried out only when the null hypothesis of unit root is rejected). The unit root test includes a linear trend only if it is significant at the 5% level. The lag length of test equations has been selected with the Schwarz information criterion. For further details see notes to Table 2.

and "High transmission capacity" identify countries that have transmission capacity below or above the median of the sample for 2017 (Source: ENTSOG Capacity Map for 2017). In countries with high transmission capacity the frequency of country pairs for which the log–price differential does not feature a unit root is higher than for country pairs with capacity below the 2017 median.

All in all, the existence of pairwise convergence seems to be positively associated with the developments of wholesale gas markets and the degree of interconnection. Empirical evidence in support of pairwise convergence is sharper for countries with high transmission capacity and well–functioning trading hubs. Of the three characteristics we have investigated, the degree of development of the trading hub seems to be the most important; in fact, for these country pairs the share of rejections of the null of a unit root in the log–price differential increases by about 10% compared with the result for the entire set of country pairs.

5 Robustness checks and additional results

5.1 Structural breaks

In Section 4.1 we have considered pairwise convergence relying on relatively standard unit root tests that have power against the alternative of trend stationarity. To take into account the possibility of structural breaks and to analyze pairwise convergence before and after key policy events, in Section 4.2 we have repeated the analysis over different time periods. In this Section we consider two alternative approaches to tackle the issue of structural breaks. To save on space tables with the results of these tests are presented in Appendix A.5 while here we comment on the evidence.

The Zivot and Andrews (2002) test. Perron (1989) showed that standard tests cannot reject the unit root null hypothesis when the true data generating process is that of a stationary series that fluctuates around a time trend with a one–time structural break that changes its slope and/or its intercept. The author proposed a test where the null hypothesis is that of a series with a unit root, while the alternative hypothesis is that the series is stationary around a broken trend. In this case, the date of the structural break is exogenously determined and the researcher has to pick an observation that identifies the time period when the trend changes intercept, slope or both. Zivot and Andrews (2002) showed that this procedures can be improved if the break date is estimated from the data. We implement this second approach and test the null hypothesis that a log–price differential can be approximated by a unit root process with drift, against the alternative hypothesis that the series is stationary around a broken trend. The test, shown in Table A3 of Appendix A.5, identifies the second semester of 2001 as a breakpoint. Interestingly, this corresponds to our analysis of convergence in Table 3, where we considered the pre– and post–Euro introduction periods. Results of the Zivot–Andrews test confirm that pairwise convergence is supported by the data. In fact, the null hypothesis is rejected a large number of times relative to the nominal size of the test.

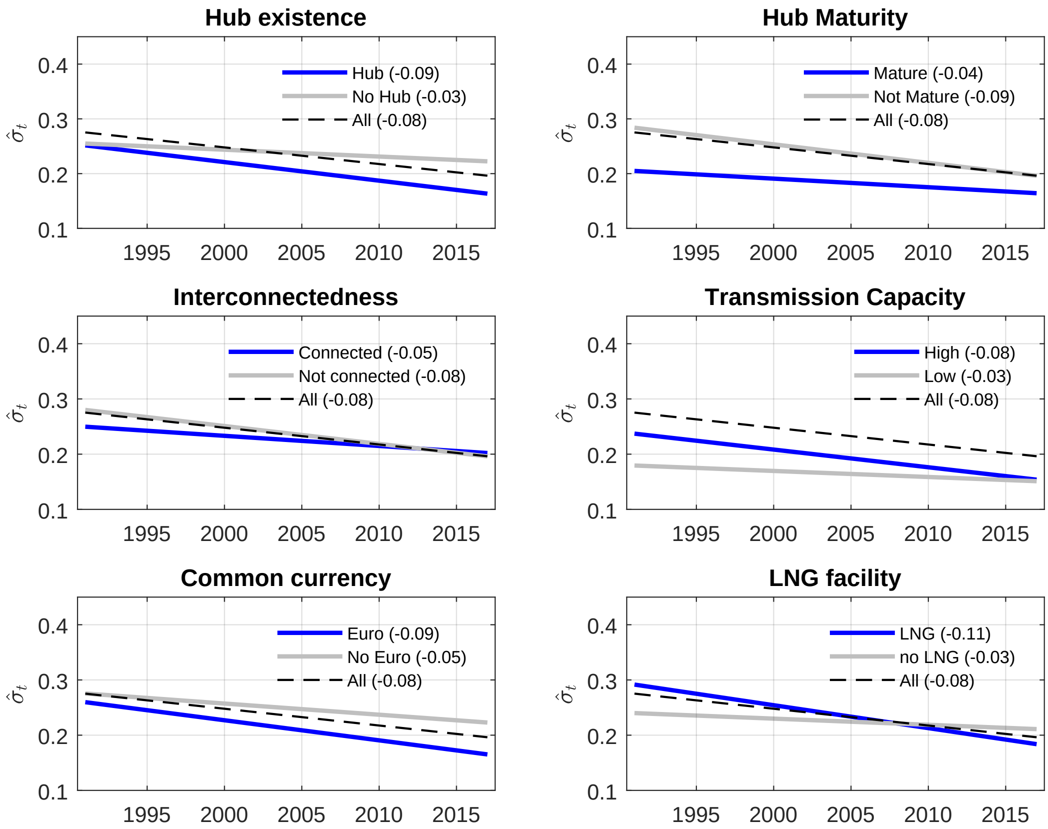
The KSS test. An alternative way of controlling for possible structural breaks in the analysis of pairwise convergence is to consider the test of Kapetanios et al. (2003) (KSS, henceforth). These authors developed a test for the null hypothesis of a unit root that has power against the alternative that the log–price differential is generated by a smooth transition autoregressive model. Pesaran et al. (2009) pointed out that the KSS test has also power against a three–regime threshold alternative. This feature of the KSS test is relevant for our analysis that is based on a relatively small sample, thus preventing us from using unit root tests that allow to accommodate more than one structural break. Results in Table A4 of Appendix A.5 show that our results are robust also when considering this test.

5.2 σ-convergence

σ–convergence is based on the idea that the cross–section variation of natural gas prices decreases over time, as we would expect from two series that converge. As shown in Appendix A.6 the notion of pairwise convergence and that of σ–convergence are intertwined. We investigate σ–convergence for different groups of countries. In addition to the classification criteria adopted in Section 4.3 (i.e. transmission capacity, existence and maturity of the trading hub), we also group countries depending on whether they share the same currency, on whether they are interconnected or not and on whether they have a Liquefied Natural Gas (LNG) regassification facility. Countries with LNG facilities were identified with circles in Figure 1 (see also Table A2).

Figure 4 shows that the cross–section standard deviations tend to decrease over time, which supports the existence of σ–convergence. Moreover, in most cases a more developed wholesale gas market is associated with a lower standard deviation. When assessing σ–convergence we are not interested only in the level of the standard deviation, but also in the slope of the trends that in Figure 4 are used to approximate their dynamics. Steeper negative trends are associated with countries pairs that have high transmission capacity, share a common currency, have a trading hubs or operate LNG terminals. Table A5 in Appendix A.6 provides statistical tests supporting these qualitative results.

Figure 4: σ–convergence for log–price differentials: 1991:h1-2017:h1



Notes: the figure tracks the dynamics of the cross–sectional standard deviations for different groups of countries. "All" means all countries, "No Hub" ("Hub") denotes countries without (with) a gas hub, "Not Mature" that stands for "Poor, Inactive or No hub" ("Mature", that is "Mature or Active hubs") indicates the classification of countries based on hub maturity (see EFET, 2016; Heather and Petrovich, 2017), "Low" ("High") indicates countries that have transmission capacity below (above) the 2017 median transmission capacity (source: 2017 ENTSOG Capacity map dataset:`https://www.entsog.eu`). "Euro" ("No Euro") indicates countries with (without) common currency. Countries with (without) LNG regassification terminals are denotes as "LNG" ("no LNG"). The figure shows cross–sectional standard deviations fitted with a linear time trend. The numerical value in the legend is the estimated trend slope.

5.3 Relative convergence

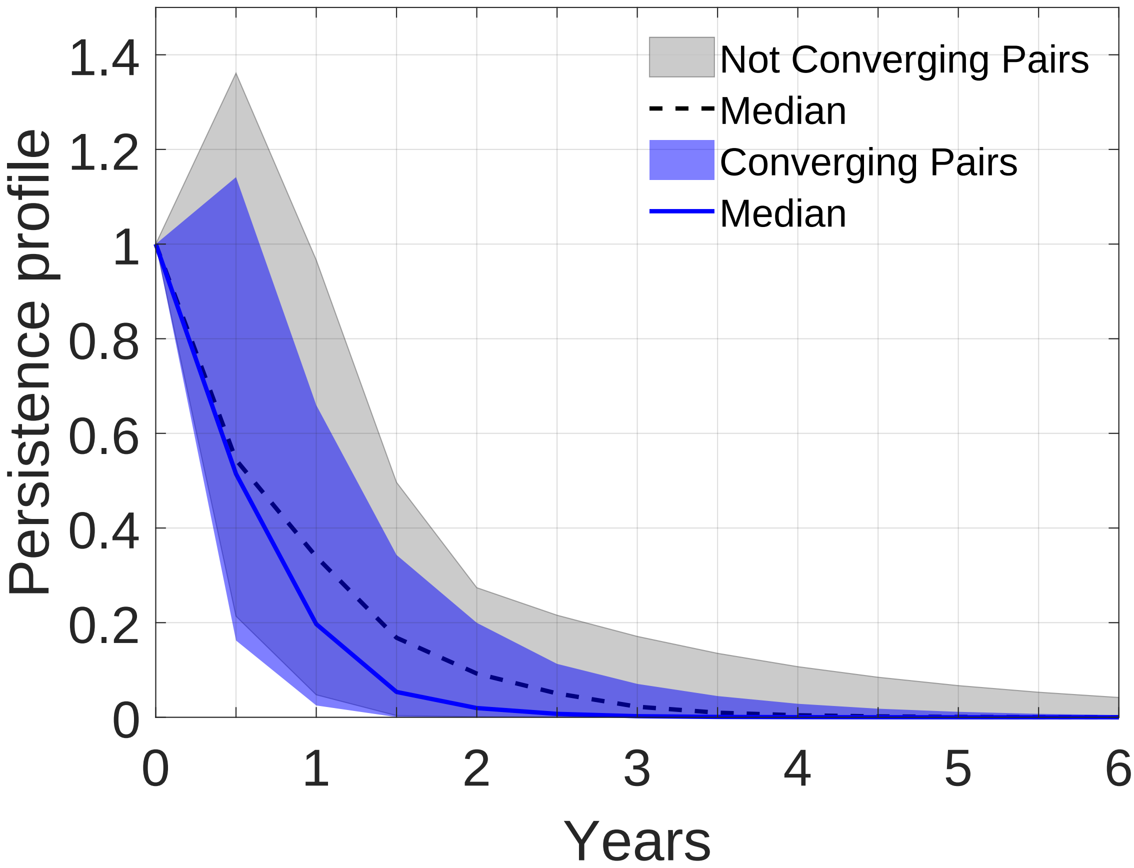
Phillips and Sul (2007) proposed a very flexible test of "relative convergence" that is valid under a less restrictive set of assumptions than those needed to satisfy the notion of pairwise convergence (see also Fischer, 2012). Relative convergence means that two series share the same stochastic or deterministic trend in the long–run and hence their ratio will eventually converge to unity. Interestingly, this notion of convergence allows for periods of transitional divergence: what matters is that in the long–run the two series converge to the common trend. Moreover, this approach can be used to investigate whether the convergence process involves the log–level of prices or their growth rates. For our data we cannot reject the null hypothesis of relative convergence. Moreover, the test provides empirical evidence that European gas prices feature growth rate convergence, but not log–level convergence. Details on the implementation of the test are provided in Appendix A.7.

5.4 Persistence profiles

While the notion of convergence has to do with the long–run behavior of prices, it is useful to provide policy makers with some evidence about its short–run implications. This can be achieved focusing on the speed with which natural gas prices return to equilibrium after a shock. To that end, we estimate persistence profiles (PPs) of log–prices for each pair of countries. PPs, popularized by Lee and Pesaran (1993) and Pesaran and Shin (1996), allow to trace the effect of a shock to one or more cointegrating relations through time. In the case of two series that have a unit root, but are not cointegrated the PP would never converge to zero, while in the case of cointegrated log-prices the effect of a shock is transitory and would eventually die out. To assess the impact of price convergence on the speed of adjustment to equilibrium, we aggregate PPs into two groups: those for country pairs with and without converging natural gas prices. Here, a pair of log-prices is defined as "converging" if the null hypothesis of unit root is rejected based on the DF–GLS test of Elliott et al. (1996) at the 10% (i.e. bottom line of column 7 in Panel (b) of Table 2; further details are provided in Appendix A.8).

Figure 5 shows estimated PPs for all pairs of converging and non–converging log–prices. Two shaded areas are plotted, one for each group of prices; these bands are bounded by the maximum and the minimum PPs and contain the median PP for the group. PPs are normalized so that the system–wide shock has a unit effect on impact. We can see that while both bands shrink toward zero, PPs for converging country pairs tend to do so much earlier. More precisely, while for converging pairs after four years there is basically no more sign of the shock, for non–converging countries PPs are still bounded between zero and 0.1, meaning that, for some prices 10% of the shock has yet to disappear. Focusing on a shorter horizon and looking at the median PP, we see that after a year 80.3% of the adjustment process

Figure 5: Persistence profiles: 1991:h1-2017:h1



Notes: the figure shows the persistence profiles estimated using a bivariate Vector Error Correction Models (VECM) of order 1 for each log-price pair. The figure distinguishes pairs for which there is evidence of pairwise convergence from those that do not display a converging behavior. In both cases we report the whole distribution of persistence profiles and their median. A pair of log–prices is converging if the null hypothesis of unit root is rejected based on the DF–GLS test of Elliott et al. (1996) at the 10%. The unit root test includes a linear trend only if it is significant at the 5% level. The lag length of the tests has been selected with the Schwarz information criterion. For further details see notes to Table 2.

for converging pairs has already been completed. This percentage is only 66.1% for non–converging prices. All in all, PPs in Figure 5 convey a very clear message: long–run price convergence helps to restore equilibrium in natural gas price markets and this has short–run policy implications. In fact, the speed of adjustment ultimately affects the welfare of citizens that have to pay higher bills for a longer time period. If, say, a supply–side oil price shock originating in a producing country hits the global economy and natural gas contracts are linked to the price of crude oil, in countries with converging prices its effects will die out earlier.

6 Discussion & Conclusions

This paper has investigated the convergence of natural gas prices in fourteen European countries. We have focused on prices paid by "medium–sized" industrial consumers over the period 1991-2017. Our empirical analysis was based on the notion of pairwise convergence that requires less restrictive hypotheses than other convergence concepts used in the literature. In addition, the chosen methodology does not require to select a benchmark price and can be applied to samples of any dimension. On the contrary, methods based on cointegration tests are not well suited in settings where the cross-sectional dimension is large.

We have found that there is evidence of pairwise price convergence and that this process is associated with key characteristics of the gas market, such as the existence and the maturity degree of development of trading hubs, as well as the degree of interconnection. This result carries over to the notion of σ–convergence and is robust to a number of changes in the implementation of the tests.

Price convergence across European gas markets is thus more likely to occur when each national gas system delivers reliable signals of its state, a feature that comes together with the establishment and maturity of gas hubs. Moreover, sufficient interconnection among national gas systems is needed, requiring to remove physical and contractual barriers to trade and arbitrage. In a perfectly interconnected European system the cross-countries arbitrage opportunities would be easily exploited by operators, pushing towards a single European price. Hence, the relevant issue refers to which frictions may prevent such super–national adjustment to take place.

Firstly, there is an issue of physical transmission capacity across countries, that may limit the ability to trade across markets and maintain price differentials. These frictions are more frequently temporary ones, due to large supply or demand shocks that would require a cross-country trade larger than transmission capacity. This occurred, for instance, in September 2016 when the interconnector between the UK and Belgium broke down, or recurrently in the interconnection between the Austrian and the German systems during the summer. The physical issue calls for infrastructural projects to increase the capacity and remove the bottlenecks.

Secondly, since transmission capacity is ruled by contracts, these latter may become another source of frictions that prevent cross-country arbitrage and price alignments. This is the case with the connection of the Italian system with the north-west European area through the Transitgas pipelines, although a reform in the congestion management procedures improved the performance from the second half of 2016. Similar issues arise in the Spanish and Polish markets. Contractual issues require a regulatory intervention to remove restrictive clauses and promote an efficient congestion management. We can notice that if interconnection is extremely efficient, even small and not very liquid national gas hubs may deliver converging prices taking advantage of the large liquidity of the gas systems they are interconnected with. The example of the infant Czech hub very well interconnected with the north-west area of the Dutch and German systems is a good example of this "substitution potential" between internal liquidity and international interconnection.

Aknowledgments

This paper was part of IEFE-Bocconi research program on "Energy Markets and Other Network Industries". We are grateful to Lutz Kilian for insightful comments on a previous version of the paper. We also thanks participants to the 1st Paris International Conference on the Economics of Natural Gas held in June 2017 and to the 6th International Symposium on Environment and Energy Finance Issues held in Paris in May 2018.

References

Apergis, N., Bowden, N., and Payne, J. E. (2015). Downstream integration of natural gas prices across US states: Evidence from deregulation regime shifts. *Energy Economics*, 49:82–92.

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

Asche, F., Oglend, A., and Osmundsen, P. (2017). Modeling UK natural gas prices when gas prices periodically decouple from the oil price. *Energy Journal*, 38(2):131 – 148.

Asche, F., Osmundsen, P., and Tveteras, R. (2002). European market integration for gas? Volume flexibility and political risk. *Energy Economics*, 24(3):249 – 265.

Auping, W., Pruyt, E., de Jong, S., and Kwakkel, J. (2016). The geopolitical impact of the shale revolution: Exploring consequences on energy prices and rentier states. *Energy Policy*, 98:390–399.

Bastianin, A., Castelnovo, P., and Florio, M. (2018). Evaluating regulatory reform of network industries: a survey of empirical models based on categorical proxies. *Utilities Policy*, 55:115–128.

Bernard, A. B. and Durlauf, S. N. (1995). Convergence in international output. *Journal of Applied Econometrics*, 10(2):97–108.

Bernard, A. B. and Durlauf, S. N. (1996). Interpreting tests of the convergence hypothesis. *Journal of Econometrics*, 71(1-2):161–173.

Brau, R., Doronzo, R., Fiorio, C. V., and Florio, M. (2010). EU gas industry reforms and consumers' prices. *The Energy Journal*, 34(4):167–182.

Bunzel, H. and Vogelsang, T. J. (2005). Powerful trend function tests that are robust to strong serial correlation, with an application to the Prebisch–Singer hypothesis. *Journal of Business & Economic Statistics*, 23(4):381–394.

Cavanagh, C. L., Elliott, G., and Stock, J. H. (1995). Inference in models with nearly integrated regressors. *Econometric Theory*, 11(5):1131–1147.

Cheung, Y.-W. and Lai, K. S. (1998). Parity reversion in real exchange rates during the post–Bretton Woods period. *Journal of International Money and Finance*, 17(4):597–614.

Cheung, Y.-W. and Lai, K. S. (2009). Practitioners Corner: Lag order and critical values of a modified Dickey–Fuller test. *Oxford Bulletin of Economics and Statistics*, 57(3):411–419.

Cremer, H. and Laffont, J. J. (2002). Competition in gas markets. *European Economic Review*, 46:928–935.

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

De Vany, A. and Walls, W. (1993). Pipeline access and market integration in the natural gas industry: Evidence from cointegration tests. *Energy Journal*, 14(4):1–20.

del Valle, A., Dueñas, P., Wogrin, S., and Reneses, J. (2017). A fundamental analysis on the implementation and development of virtual natural gas hubs. *Energy Economics*, 67:520–532.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a):427–431.

EFET (2016). Review of gas hub assessment. Available online at (last accessed November, 2017): `https://www.efet.org`.

EIA (2017). International Energy Outlook 2017. U.S. Energy Information Administration.

Elliott, G. (1998). On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica*, 66(1):149–158.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836.

Fischer, C. (2012). Price convergence in the EMU? Evidence from micro data. *European Economic Review*, 56(4):757–776.

Growitsch, C., Stronzik, M., and Nepal, R. (2015). Price convergence and information efficiency in german natural gas markets. *German Economic Review*, 16(1):87–103.

Heather, P. (2012). The evolution of European traded gas hubs. OIES paper NG 104, Oxford Institute for Energy Studies.

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

Hodrick, R. J. and Prescott, E. C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Hulshof, D., van der Maat, J.-P., and Mulder, M. (2016). Market fundamentals, competition and natural-gas prices. *Energy Policy*, 94:480 – 491.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian Vector Autoregressive models. *Econometrica*, 59(6):1551–1580.

Kapetanios, G., Shin, Y., and Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, 112(2):359–379.

Kilian, L. (2017). The impact of the fracking boom on Arab oil producers. *Energy Journal*, 38(6):137–160.

King, M. and Cuc, M. (1996). Price convergence in North American natural gas spot markets. *The Energy Journal*, 17(2):17–42.

Koske, I., Wanner, I., Bitetti, R., and Barbiero, O. (2015). The 2013 update of the OECD's database on product market regulation. Working Paper 1200, OECD Economics Department.

Koster, H. R. and van Ommeren, J. (2015). A shaky business: Natural gas extraction, earthquakes and house prices. *European Economic Review*, 80:120–139.

Lee, K. C. and Pesaran, M. H. (1993). Persistence profiles and business cycle fluctuations in a disaggregated model of UK output growth. *Ricerche Economiche*, 47(3):293–322.

Leybourne, S., Kim, T.-H., and Newbold, P. (2005). Examination of some more powerful modifications of the Dickey–Fuller test. *Journal of Time Series Analysis*, 26(3):355–369.

Li, R., Joyeux, R., and Ripple, R. D. (2014). International natural gas market integration. *The Energy Journal*, 35(4):159–179.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11(6):601–618.

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

Neumann, A., Siliverstovs, B., and 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.

Nick, S. and Thoenes, S. (2014). What drives natural gas prices? — A structural VAR approach. *Energy Economics*, 45:517 – 527.

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

Park, H. J. and Fuller, W. A. (1995). Alternative estimators and unit root tests for the autoregressive process. *Journal of Time Series Analysis*, 16(4):415–429.

Perron, P. (1989). The Great Crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57(6):1361–1401.

Pesaran, M. H. (2007). A pair–wise approach to testing for output and growth convergence. *Journal of Econometrics*, 138(1):312–355.

Pesaran, M. H. and Shin, Y. (1996). Cointegration and speed of convergence to equilibrium. *Journal of Econometrics*, 71(1-2):117–143.

Pesaran, M. H., Smith, R. P., Yamagata, T., and Hvozdyk, L. (2009). Pairwise tests of purchasing power parity. *Econometric Reviews*, 28(6):495–521.

Phillips, P. C. B. and Sul, D. (2007). Transition modeling and econometric convergence tests. *Econometrica*, 75(6):1771–1855.

Ravn, M. O. and Uhlig, H. (2002). On adjusting the Hodrick–Prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–376.

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

Robinson, T. (2007). Have European gas prices converged? *Energy Policy*, 35(4):2347–2351.

Saussay, A. (2018). Can the US shale revolution be duplicated in continental Europe? An economic analysis of european shale gas resources. *Energy Economics*, 69:295 – 306.

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

Siliverstovs, B., L'Hégaret, G., Neumann, A., and 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.

Stock, J. and Watson, M. (2012). Disentangling the channels of the 2007-2009 recession. *Brookings Papers on Economic Activity*, Spring 2012:81–135.

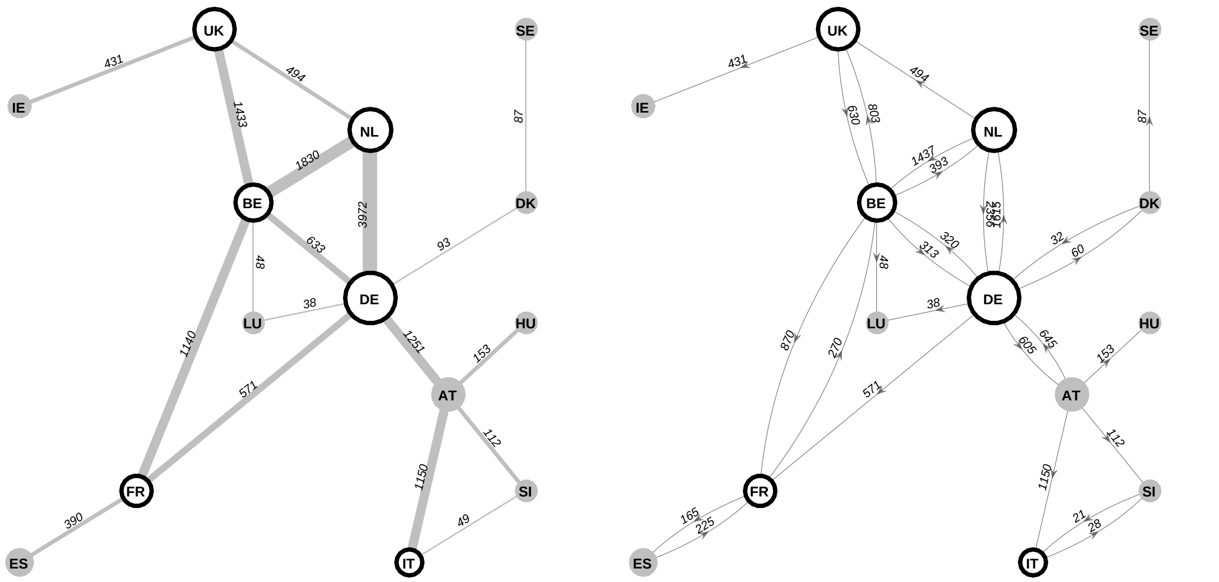
Zivot, E. and Andrews, D. W. K. (2002). Further evidence on the Great Crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 20(1):25–44.

A Appendix

A.1 Natural gas markets and technical physical capacity

Figure A1 displays an undirected and a directed network graph of the natural gas markets we are studying.[7] In these graphs the size of nodes is proportional to the country's technical physical capacity, given by the total of within EU inflows and outflows, export to non–EU countries and technical physical capacity at LNG terminals. In panel (a) the thickness of the edges connecting the nodes is proportional to the technical physical capacity between two countries, while the arrows in panel (b) indicate the direction of the gas flow. It is important to note that we consider only the capacity of the interconnections between the countries in our sample.

Figure A1: Undirected & directed network graph of the EU natural gas market in 2017



(a) Undirected graph (b) Directed graph

Notes: Size of nodes is proportional to the country's technical physical capacity (GWh/d) given by the total of within EU inflows and outflows, export to non–EU countries and technical physical capacity at LNG terminals. Size of edges is proportional to technical physical capacity between the two countries. Countries with mature and active hubs highlighted with a bold black line. Authors' elaborations using the 2017 ENTSOG capacity map dataset (`https://www.entsog.eu`).

[7]Figure A1 relies on the 2017 ENTSOG capacity map dataset (`https://www.entsog.eu`).

A.2 The OECD ETCR database

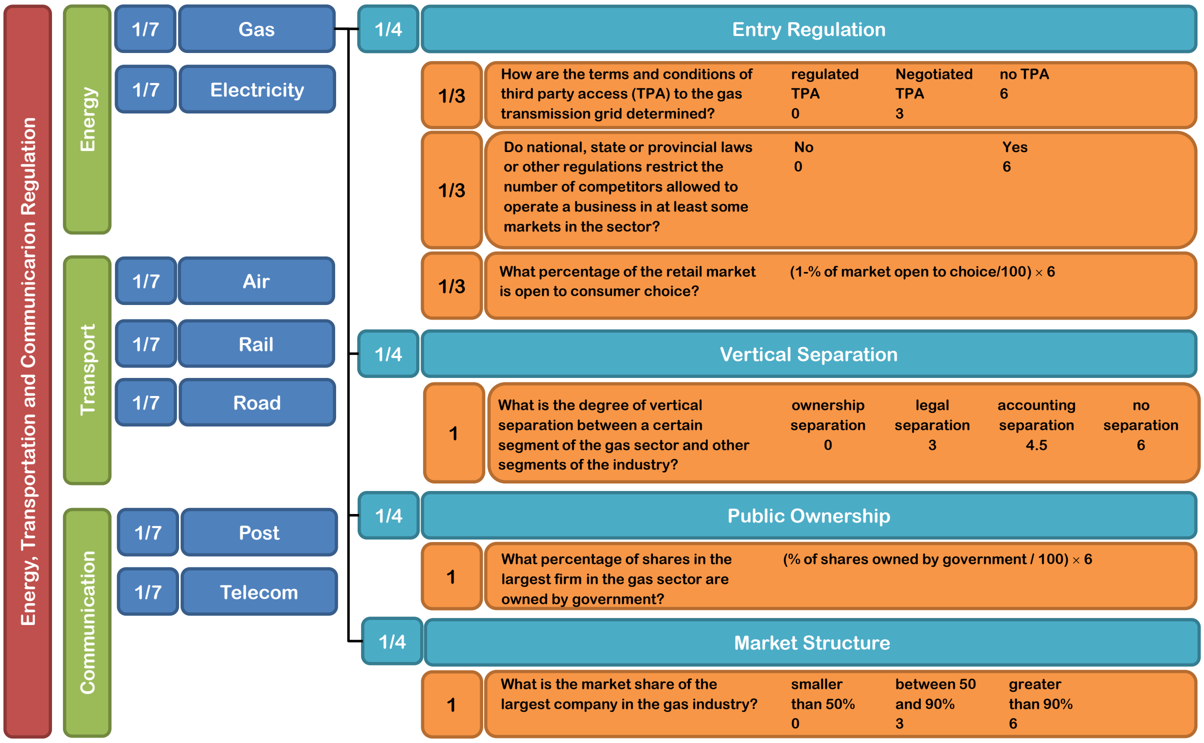
The OECD indicators of regulation in energy, transport, and communications (ETCR) collect survey information about regulatory structures and policies for OECD and some non–OECD countries (see Koske et al., 2015, for details). All answers are normalized in a range from zero to six, where values near zero indicate fewer restrictions to competition (see Bastianin et al., 2018, for a discussion of categorical proxies of reform). As shown in Figure A2, the ETCR index is available at different levels of aggregation. The Product Market Regulation (PMR) index in column 6 of Table 1 aggregates with equal weights the sub–indicators for seven network industries: natural gas, electricity, air, rail, road transport, post and telecommunications. For each sector, up to four dimensions of regulatory policy are analyzed: entry regulation, public ownership, vertical integration and market regulation. These sub-indexes for gas markets are shown in Table A1, while Figure 3 in the main text shows the aggregate index for natural gas.

Table A1: Descriptive statistics for the European natural gas markets in 2017

		ETCR score				
(1)	(2)	(3)	(4)	(5)	(6)	(7)
iso	country	Gas	ER	PO	VI	MS
AT	Austria	2.2	0.0	2.8	4.7	1.5
BE	Belgium	1.7	0.0	2.2	4.5	0.0
DE	Germany	1.2	0.0	0.0	4.7	0.0
DK	Denmark	2.6	0.0	4.5	4.5	1.5
ES	Spain	1.1	0.0	0.1	3.0	1.5
FR	France	2.5	0.0	2.4	4.7	3.0
HU	Hungary	1.7	1.0	0.6	3.2	2.3
IE	Ireland	3.0	0.0	5.8	4.5	1.5
IT	Italy	1.9	0.0	1.8	4.9	0.8
LU	Luxembourg 2.6	0.0	2.8	4.7	3.0	
NL	Netherlands	2.3	0.5	3.5	4.5	0.8
SE	Sweden	1.7	0.0	0.0	3.8	3.0
SI	Slovenia	2.8	0.0	3.3	4.9	3.0
UK	UK	0.0	0.0	0.0	0.0	0.0

Notes: Columns 3–7 report ETCR scores sourced from the OECD's Energy, Transport and Communications Regulation (Koske et al., 2015). These data represent a scoring system on a scale from zero to six, where values near zero indicate fewer restrictions to competition. We report an overall score for the gas market ("Gas") that is the average of the scores for entry regulation ("ER"), public ownership ("PO"), vertical separation ("VS") and market structure ("MS").

Figure A2: Structure of the OECD's Energy, Transport and Communications Regulation (ETCR) database.

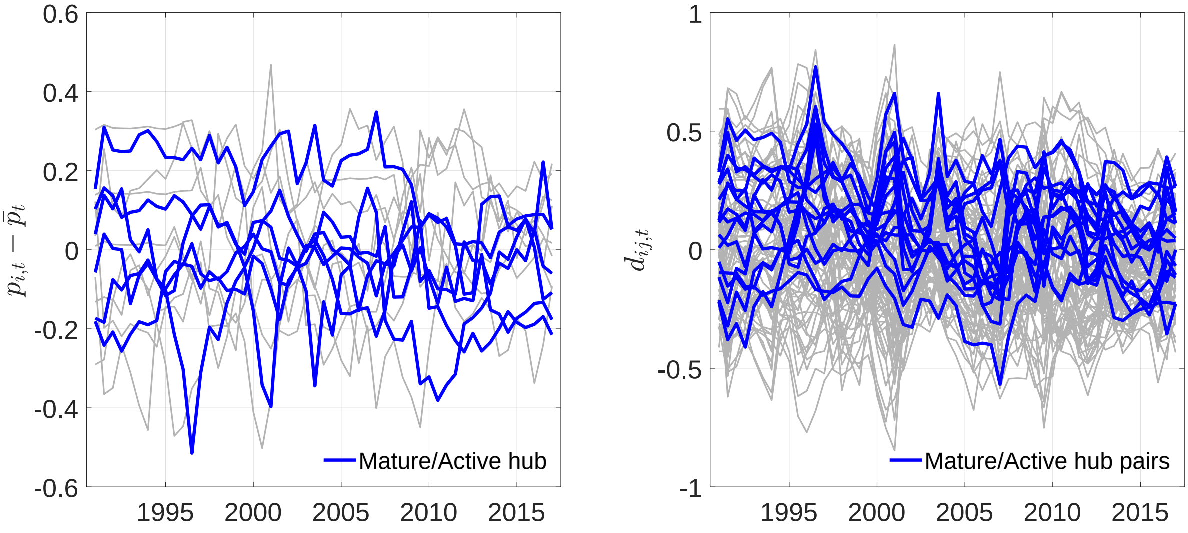


Notes: The ETCR index aggregates with equal weights indexes for seven network sectors: telecommunication, electricity, gas, post, air transport, rail transport, and road transport. For each sector, there are up to four sub-indexes that cover different dimensions of the reforms: entry regulation, public ownership, vertical integration and market regulation. We show the underlying questionnaire for the gas sector. Numbers are sector–, topic– and question–weights used for aggregation purposes. Source: (Bastianin et al., 2018).

A.3 Log-price differentials: visual inspection

Figure A3 displays log–prices as differences from the cross–country average (a) and log–price differentials ($d_{ij,t} = p_{i,t} - p_{j,t},$) (b). Recall that pairwise convergence implies that two prices converge if $d_{ij,t}$ is stationary with constant mean. With $N = 14$ countries, we have a total of $N(N-1)/2 = 91$ log–price differentials to test. Visual inspection of Figure A3(b) does not provide any clear evidence in support of, or against convergence. However, in countries with an active or mature trading hub (identified by darker lines) log–price differentials tend to be less dispersed. This is even more evident in Panel (a), where at the end of the sample the spread of the log-prices (in difference from the cross–country average) decreases, especially in countries where there is an active or mature trading hub. These facts qualitatively support

Figure A3: Log–prices & differentials: 1991:h1-2017:h1



(a) Log–price: difference from cross–country sample average

(b) Log–price differential

Notes: darker lines indicate "Mature/Active Hubs" in the classification of countries based on hub maturity (see EFET, 2016; Heather and Petrovich, 2017).

the existence of σ–convergence, namely the tendency for price differentials to become less dispersed over time (see also Section 5.2 in the main text and Section A.6 in the Appendix and .

A.4 Country groupings

Table A2: Groups of countries

(a) **Common currency**
• *Euro:* Austria, Belgium, Germany, Spain, France, Ireland, Italy, Luxembourg, Netherlands, Slovenia
(b) **Natural gas hub**
• *Countries with a trading hub:* Austria, Belgium, Germany, Denmark, Spain, France, Hungary, Ireland, Italy, Netherlands, United Kingdom
• *Mature or Active hubs:* Belgium, Germany, France, Italy, Netherlands, United Kingdom
• *Poor, Inactive or no hubs:* Austria, Denmark, Spain, Hungary, Ireland, Luxembourg, Sweden, Slovenia
(c) **Transmission capacity**
• *High:* Austria, Belgium, Germany, Spain, France, Netherlands, United Kingdom
• *Low:* Denmark, Hungary, Ireland, Italy, Luxembourg, Sweden, Slovenia
(d) **No. of interconnections**
• *1 Interconnection:* Spain, Hungary, Ireland, Sweden
• *2 Interconnections:* Denmark, Italy, Luxembourg, Slovenia
• $\geq$ *3 Interconnections:* Austria, Belgium, Germany, France, Netherlands
(e) **LNG regassification facility**
• *Countries with LNG terminals:* Belgium, Spain, France, Italy, Netherlands, United Kingdom

Notes: In panel (a) "Euro" indicates countries with common currency. Panel (b) lists countries with a trading hub and their development level. "Mature or Active hubs" and "Poor, Inactive or No hub" indicate the 2016 classification of countries that is based on hub maturity (see Section 2.1, EFET, 2016; Heather and Petrovich, 2017). In Panel (c) "High" ("Low") indicates countries that have transmission capacity above (below) the 2017 median transmission capacity for the markets in our sample. Panel (d) divides gas markets according to the number of interconnections between the countries in our sample. Panel (e) lists countries with one or more LNG regassification terminal. Information in Panel (c-e) is based on: the 2017 ENTSOG Capacity map dataset:`https://www.entsog.eu`.

A.5 Structural breaks

Table A3: Pairwise convergence based on the Zivot and Andrews (2002) test: 1991:h1–2017:h1

(1) Model	(2) 5%	(3) 10%	(4) break date
Break in intercept	39.56	47.25	2001:$h2$
Break in intercept & trend	37.36	47.25	2001:$h2$

Notes: columns (2-3) show the percentage of the 91 log–price differentials ($d_{ij,t}$) for which the null hypothesis of unit root is rejected relying on the Zivot and Andrews (2002) test ($\bar{Z}_{NT}$). The alternative hypothesis of the test is that $d_{ij,t}$ is stationary process with a break in the intercept or in the intercept and trend, see column (1). Convergence between prices is supported by the data when $\bar{Z}_{NT}$ is large relative to the significance level of the unit root test — 5% or 10% in this table. The lag length of test equations has been selected with the Akaike information criterion. The maximum lag order is set equal to 4 that corresponds to two years. The test is performed discarding 20% of the observations at the beginning and at the end of the sample. The date in column (5) is the date of the break that is identified most often by the test.

Table A4: Pairwise convergence based on the Kapetanios et al. (2003) test: 1991:h1–2017:h1

	5%				10%			
	(1) AIC	(2) %t	(3) SIC	(4) %t	(5) AIC	(6) %t	(7) SIC	(8) %t
const	40.7	35.1	39.6	38.9	52.7	33.3	47.3	34.9
const/trend	46.2	31.0	42.9	35.9	56.0	31.4	49.5	33.3

Notes: columns (1, 3, 5, 7) show the percentage of the 91 log–price differentials ($d_{ij,t}$) for which the null hypothesis of unit root is rejected relying on the Kapetanios et al. (2003) test($\bar{Z}_{NT}$). The alternative hypothesis of the test is that $d_{ij,t}$ follows a nonlinear, but stationary process. In the case of rejection of the null hypothesis (i.e. for stationary log–price differentials), columns (2, 4, 6, 8) show the percent of log–price differentials for which the hypothesis of a non–significant trend is not rejected. Student tests of the significance of the linear trend are conducted at the 5% significance level in a regression of $d_{ij,t}$ on a constant and a linear trend (the test is carried out only when the null hypothesis of unit root is rejected). Convergence between prices is supported by the data when $\bar{Z}_{NT}$ is large relative to the significance level of the unit root test — 5% or 10% in this table — and the number of trend stationary series relatively low (i.e. the % in the "%t" column is high). We present three cases for the deterministic component of the Kapetanios et al. (2003) test: "const" indicates that we included only the intercept; "trend" indicates that we included the intercept and a linear trend; "const/trend" indicates that we included a linear trend only if it is significant at the 5% level. The lag length of test equations has been selected either with the Akaike or with the Schwarz information criterion, denoted as AIC and SIC, respectively. The maximum lag order is set equal to 4 that corresponds to two years. Critical values are provided by Kapetanios et al. (2003).

A.6 σ-convergence: conceptual aspects and empirical results

From pairwise to σ − convergence. Pesaran (2007) showed that starting from squared log–price differentials it is possible to construct an average measure of convergence/divergence that can be used to investigate whether the cross–section variance of prices decreases over time. In particular, the cross–section standard deviation of log–prices is proportional to the squares of the log-price differentials, $d_{ij,t}$, averaged across country pairs. That is:

$$
\begin{aligned}
D_t^2 &= \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (d_{ij,t})^2 \\
&= 2 \times \left\{ \frac{\sum_i^N (p_{i,t} - \bar{p}_t)^2}{N-1} \right\} = 2\hat{\sigma}_t^2
\end{aligned} \tag{1}
$$

where $\bar{p}_t = N^{-1} \sum_{i=1}^{N} p_{i,t}$ and D_t is proportional to the cross–section standard deviation of prices, $\hat{\sigma}_t$. The notion of σ–convergence can thus be investigated tracking $\hat{\sigma}_t$ or other measures of dispersion over time.

Further results. Table A5 provides statistical tests supporting the qualitative results in Section 5.2 of the main text. These tests are based on OLS estimation of the following model: $\hat{\sigma}_{(g)t} = \alpha_g + \beta_g trend + u_{(g)t}$ where g denotes a given set of countries. The third column of Table A5 highlights that countries with trading hubs, LNG facilities and high transmission capacity feature the steepest negative trend slopes. Moreover, columns 4-7 report statistical tests for the null hypothesis of "no σ–convergence" ($H_0 : \beta_g \geq 0$) against the alternative hypothesis of σ–convergence ($H_0 : \beta_g < 0$). Although the t–statistic in the fourth column is based on heteroskedasticity and autocorrelation consistent standard errors, when $u_{(g)t}$ has strong serial correlation or a unit root the test tends to over–reject the null (Bunzel and Vogelsang, 2005). For this reason, in columns 5-7 we also present the "Dan–J" statistic due to Bunzel and Vogelsang (2005) that includes an adjustment factor to control for the the over–rejection problem.[8]

Independently of the test being used, the null of "no σ–convergence" cannot be rejected for country pairs with low transmission capacity, without trading hubs or without LNG facilities. To further investigate whether the trend slope differs between pairs of countries in different groups, the last column of Table A5 proposes a test of the null hypothesis of common trend slopes.[9] Overall, we confirm the qualitative evidence in Figure 4: not only

[8]The "Dan–J" test can be written as follows: $t_{Dan} = \left[\frac{\hat{\beta}-\beta}{se(\hat{\beta})} \right] exp(-bJ)$ where $se\left(\hat{\beta}\right)$ are robust standard errors based on the Daniell kernel, b is a pre–specified constant that depends on the significance level of the test and $exp(-bJ)$ is a correction factor used to take into account that $u_{(g)t}$ might have strong serial correlation or a unit root. See Bunzel and Vogelsang (2005) for details.

[9]This test is implemented estimating a bivariate Seemingly Unrelated Regression model where the cross-

Table A5: σ–convergence for different country groups: 1991:$h1$–2017:$h1$

				Dan-J stat			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
g	Group	$\hat{\beta}_g$	t-stat	10%	5%	1%	$\beta_g = \beta_{g+1}$
0	All	-0.079	-3.987***	-2.418^c	-2.125^b	-1.405	-
1	Hub	-0.088	-5.469***	-2.962^c	-2.619^b	-1.765	0.0167
2	No hub	-0.032	-0.789	-0.597	-0.494	-0.274	-
3	Mature or Active hubs	-0.041	-1.959**	-1.318	-1.197	-0.880	0.0239
4	Poor, Inactive or No hub	-0.088	-4.157***	-2.549^c	-2.222^b	-1.432	-
5	Connected	-0.048	-2.360**	-0.980	-0.739	-0.300	0.0387
6	not Connected	-0.084	-3.620***	-2.258^c	-1.956^b	-1.237	-
7	Euro	-0.095	-4.526***	-3.182^c	-2.797^b	-1.853	0.2175
8	No Euro	-0.053	-1.233	-0.796	-0.656	-0.355	-
9	High transmission capacity	-0.083	-4.324***	-2.312^c	-2.017^b	-1.304	0.0444
10	Low transmission capacity	-0.028	-1.119	-0.679	-0.551	-0.281	-
11	LNG facility	-0.108	-5.084***	-2.891^c	-2.453^b	-1.450	0.0000
12	no LNG facility	-0.029	-1.262	-0.811	-0.705	-0.450	-

Notes: Column (2) shows which countries have been used in the computation of the σ–convergence test, where: "All" means all countries, "No Hub" ("Hub") denotes countries without (with) a trading hub, "Poor, Inactive or No hub" ("Mature or Active hubs") indicates the classification countries based on hub maturity discussed in Section 2.1 (see EFET, 2016; Heather and Petrovich, 2017), "Low transmission capacity" ("High transmission capacity") indicates countries that have transmission capacity below (above) the 2017 median transmission capacity (source: 2017 ENTSOG Capacity map dataset:`https://www.entsog.eu`). Countries with (without) LNG regassification terminals are denotes as "LNG" ("no LNG"). Column (3) shows the estimated trend slope from the following models: $\hat{\sigma}_{(g)t} = \alpha_g + \beta_g trend + u_{(g)t}$ for $g = 0, \dots, 12$. Column (4) reports the t-statistic for the null hypothesis of "no σ–convergence" ($H_0 : \beta_g \geq 0$) against the alternative hypothesis "σ–convergence" ($H_0 : \beta_g < 0$). The t-statistic is based on heteroskedasticity and autocorrelation consistent standard errors. Columns (5-7) report the Bunzel and Vogelsang (2005) test. The "Dan-J" statistic is used to correct the standard t-tests used to gauge the statistical significance of the trend slope when the error term of the regression is highly persistent and possibly integrated of order one. The null and alternative hypothesis are the same as those for the standard t-tests in Column (4); since the test is a function of a pre–specified constant that depends the significance level of the tests, we report three values of the test in columns (5-7), where the superscripts a, b, c (or ***, **, *) denote the rejection of the null hypothesis of no σ–convergence at the at the 1%, 5% and 10% levels of significance. Column (8) reports a two-sided test for the equality of trend slopes. The null hypothesis is: $H_0 : \beta_g - \beta_{g+1} = 0$ for $g = 1, 3, 5, 7, 9, 11$. These tests are based on the estimation of bivariate Seemingly Unrelated Regression models where common regressors are the intercept and the trend and the dependent variables are the standard deviations of price differentials ($\hat{\sigma}_{(g)t}, \hat{\sigma}_{(g+1)t}$) for countries in each category g for $g = 1, 3, 5, 7, 9, 11$ (e.g. countries with and without a trading hub).

countries with trading hubs, LNG facilities and high transmission capacity are associated with the steepest negative slope, but for these countries estimated trends are statistically different from those for the complementary group of country pairs. For instance, focusing on the existence of a trading hub, we can see from columns 4-7 that there is evidence of a statistically significant σ–convergence pattern only for countries with a trading hub. In fact, for countries without a trading hub we cannot reject the null of no convergence. Moreover, the last column of the table shows that these two groups of countries do not have a common trend slope. Interestingly, when analyzing countries with and without a common currency, we discover that while there is evidence of statistically significant negative trend slope only

section standard deviations of two groups of countries are regressed on a trend and an intercept.

for country pairs where the Euro is the common currency, the null of a common trend slope cannot be rejected. This fact, might suggest that σ–convergence is chiefly affected by gas market characteristics, rather than by other macroeconomic factors, such as the existence of common monetary policy.

A.7 Relative convergence

The starting point of the Phillips and Sul (2007) test for relative convergence is a non–linear panel model for prices:

$$p_{it} = \delta_{it}\mu_t \tag{2}$$

where μ_t is common stochastic or deterministic trend and δ_{it} is country specific slope that captures a time–varying factor loading coefficient attached to μ_t. The role of δ_{it} is that of a vector of weights that measure the distance between the price of natural gas in country i and the common trend μ_t. Relative convergence is formally defined as:

$$\text{plim}_{t\to\infty}\frac{p_{it}}{p_{ij}} \to 1 \quad \text{for } i \neq j \tag{3}$$

where "plim" denotes convergence in probability. The Phillips and Sul (2007) test is based on the estimation of the following regression model:

$$\log\left(\frac{H_1}{H_t}\right) - 2\log\log t = \lambda_0 + \lambda_1 logt + u_t \tag{4}$$

where $H_t = N^{-1}\sum_{i=1}^{N}(h_{it}-1)^2$ is the sample transition distance and $h_{it} = p_{it}/\sum_{i=1}^{N} p_{it}$ is the relative transition curve for $t = [rT], [rT]+1, \ldots, T$ for $r = 0.3$. The null hypothesis of convergence $H_0 : \lambda_1 \geq 0$ against $H_1 : \lambda_1 < 0$ can be tested with a standard one–sided t–test based on heteroskedasticity and autocorrelation consistent standard errors. The magnitude of the slope coefficient of $\log t$ is also of interest. In fact, it can be shown that we have growth convergence if $0 \leq \lambda_1 < 2$ and convergence in log-level if $\lambda_1 \geq 2$.

To implement the test we first use the Hodrick–Prescott filter to smooth out business cycle components from the log–price series (see notes to Figure 2 for details on HP filtering). OLS estimation of Equation (4) yields $\hat{\lambda}_1 = 0.4620$, with a corresponding standard error of

0.0973. This implies that the t-statistic is equal to 4.7501 which is greater than the one–sided critical values at any standard level of significance: therefore, we cannot reject the null hypothesis of relative convergence. Moreover, since $0 \leq \hat{\lambda}_1 < 2$, there is evidence that European gas prices feature growth rate convergence, but not log–level convergence.

A.8 Persistence profiles: estimation and further details

Persistence profiles (PPs) are formally defined as the scaled difference between the conditional variance of M–step and $(M-1)$–step ahead forecasts. This difference is plotted as a function of the forecast horizon M. PPs are thus "variance–based" measures of the persistence of shocks to the system at different time horizons.

There are two main differences between PPs and impulse–response functions (IRF). First, IRFs trace the effect of a particular shock on a given variable, while PPs focus on the time profile of a system–wide shocks on the cointegrating relation. For this reason, interpretation of PPs does not require to orthogonalize the covariance matrix of the shocks, which is often viewed as the most critical part of structural IRF analysis.

Estimation of PPs relies on a set of bivariate Vector Error Correction Models (VECM), one for each pair of countries. In our application, one lag was sufficient to remove the serial correlation in the residuals. Including one lag in the VECM for first–differenced log–prices means that the underlying Vector Autoregressive model for the log–level of the variables is of order two. Estimation of the VECM relies on the approach due Johansen (1991).

GREEN
Centre for Geography, Resources, Environment, Energy and Networks
via Röntgen, 1
20136 Milano - Italia

www.green.unibocconi.eu