World Coal Markets: Still Weakly Integrated and Moving East

Bo Liu[a], Hélyette Geman[a,b,*]

[a] *Department of Applied Mathematics and Statistics, Johns Hopkins University, 3400 North Charles St, Baltimore, MD 21218, USA*
[b] *Department of Economics, Mathematics and Statistics, Birkbeck, University of London, Malet St, London WC1E 7HX, UK*

Abstract

In December 2016, the prominent New York Mercantile Exchange CAPP coal future contract was delisted by the CME, owner of the NYMEX. At that time, all four giant US coal miners were under the protection of Chapter 11. These events illustrate the collapse of coal consumption in the US and the loss of relevance of the Appalachian coal index whose futures were used as a hedging instrument across the world. Our goal in this paper is to revisit the problem of integration of coal markets and to exhibit through multiple perspectives, including a lead/lag analysis of pairs of major indexes, that the world coal market is moving East.

Keywords: Coal indexes, cointegration, lead and lag analysis, US miners' equities

1. Introduction

Coal continues to be the predominant fuel for electricity generation worldwide, and a main source for global energy supply. While its growth has slowed as it is losing market share to presently cheap natural gas, coal consumption is still

*Corresponding author
Email address: hgeman@hotmail.com (Hélyette Geman)

We wish to thank three anonymous referees and the Editor for their suggestions on an earlier version of the paper.

expected to grow at 0.5% per annum over the next 20 years and maintain 25% of global energy consumption in 2035 (BP Energy Outlook 2016).

Much of the energy increase in the coming two decades can be attributed to the increase of power generation as the trend towards global electrification continues, in poorer regions of the planet in particular. Coal generates today over 41% of the world's electricity and should maintain at least a 33% share through 2035, while electricity consumption should rise by as much as 43% during the same period. As for coking coal, it is a primary constituent in the production of over 70% of the world's steel. The global market for coal, both thermal and metallurgical, is currently estimated to be nearly 8 billion tonnes per year and expected to grow 12.5% by 2035, with the growth in demand centered primarily in Asia (BP Energy Outlook 2016)

In line with this numbers, our goal in this paper is dual: show, in agreement with Zaklan et al. (2012), that the world coal market continues to be weakly integrated; but is moving East from multiple perspectives.

The paper is organized as follows: Section 2 provides an overview of the global market and a summary of the literature. Section 3 presents ten major coal indexes and, using Principal Component Analysis (PCA) and cointegration analysis, exhibits that a weak integration of the market continues to prevail. Section 4 depicts the dramatic collapse of all four US coal giants' share prices, consumption numbers going forward and conducts a lead/lag analysis of six pairs of major (Eastern, Western) indexes exhibiting a clear change in favor of the Eastern indexes. Section 5 concludes.

2. World Coal Markets Outlook and Literature Overview

The world coal market has been divided for a number of years into the Atlantic and Pacific regions (see Geman, 2005; Wårell, 2005), and continues to be so, in particular because of the cost of bunker fuel compared to the value of the commodity transported.[3] The Pacific region has been playing a more important role in the physical coal trade in recent years, with the growing importance of India and China on the demand side and Indonesia on the supply side. In parallel, the shale explosion in the oil and gas industry in the United States had a fundamental impact on the American coal market, and will continue to do so independently of the government views as long as US natural gas prices remain extremely low. In this context, one of our questions is to ask whether the Pacific and Atlantic regions are still divided.

Our data for this study covers nine countries: China, Japan, Korea and Australia in the Pacific market and the US, Colombia, South Africa and the Netherlands in the Atlantic market. There are two data series for Russia—Vostochny in the Pacific region and Baltic in the Atlantic region. This list of countries includes most of the important players on both the demand and supply sides, except only India[4] and Indonesia for which data are not available. Netherlands is included in our data despite its not being a top importer or exporter any more, because the API 2 index—a CIF index including cost of insurance and freight to the ARA (Antwerp-Rotterdam-Amsterdam) region of Northwestern Europe—is the most traded coal index as far as derivatives (forwards, swaps) are concerned. Compared to the Pacific market, the Atlantic market is more dominated by derivatives and paper trades instead of spot physical trades, as discussed by Papież & Śmiech (2013).

There have been several papers in the past decade investigating the integration of the international steam coal market. In two early studies, Ekawan & Duchêne (2006) and Ekawan et al. (2006) describe the hard coal markets in the Atlantic and Pacific regions respectively, and allude to the integration between the two markets, without empirically testing the hypothesis. Wårell (2006) examines

[3] Note that the fascinating topic of coal transportation and logistics is outside the scope of this paper.

[4] As a top three coal importer in the world, India is expected to import around 160 million MT of coal in fiscal year 2016–17, or about 18% of its total demand. However, Coal India Ltd, the world's largest producer of thermal coal announced in July 2016 that it was looking to increase supply and explore avenues to start exporting in the near future.

whether the European and Japanese markets were integrated over the period 1980–2000 for both the steam coal and coking coal markets using quarterly import data. For steam coal, the result supported the hypothesis of an integrated market. In order to test market integration over time, the author also applied cointegration tests to two sub-periods, the 1980s and the 1990s and concluded that no integration could be confirmed for the 1990s. Li et al. (2010) investigated the hypothesis of a single economic market for the international steam coal industry. Using multiple methods including cointegration tests and the Kalman filter, with monthly FOB export data in six countries (Australia, China, Colombia, Indonesia, Poland and South Africa) between 1995–2007, they conclude that, in general, the international steam coal market could be seen as integrated over that time period. Zaklan et al. (2012) add the logistics perspective and analyze the integration of international steam coal trade using weekly import, export and transportation data from 2001 to 2009. They use a principal component analysis (PCA) and Johansen cointegration tests and concluded that there was a significant, yet incomplete, integration. Papież & Śmiech (2013) use instead a causality methodology to investigate the international steam coal market integration, especially dependencies between different markets. They analyzed weekly export and import data in seven markets (Indonesia, Australia, Japan, Korea, Netherlands, South Africa and Colombia) from 2002 to 2011. They concluded that the most important factor in the Pacific market was Australian coal; while in the Atlantic market, ARA and Richard Bay prices had the biggest influence on other prices—results which confirm the role of Australia and South Africa as a major producing country in each region. In a recent study, Papież & Śmiech (2015) conducted rolling cointegration analysis on a weekly data set of six markets (ARA, South Africa, Colombia, Russia Baltic, Indonesia, and Australia) from 2001 to 2014 and found that these markets were integrated, especially during the period when freight costs were low.

A global oil market has been identified very early on (Sauer, 1994), but the same result has not been reported for other fossil fuels. Doane & Spulber (1994) studied US natural gas in the 1980s and found a national market integrated by open access transportation. They started a literature on market integration for natural gas in the US. Siliverstovs et al. (2005) were the first to investigate the integration between European and US gas markets. Geman & Liu (2015) analyzed both spot prices and forward curves for the US Henry Hub and UK National Balancing Point natural gas indexes and concluded to the absence of convergence of these two markets despite the general decline in gas prices.

We apply in Section 3 the methods of PCA and cointegration that have been previously used, but to a richer dataset. The ten countries we cover include emerging players in global coal trade markets such as China, as well as big exporters which provide coal to both Asian and European markets, namely Russia and the US. These countries are often omitted in existing literature.

3. Data and Statistical Analysis

To have a clearer picture of the relative contributions of the Pacific and Atlantic regions to the world coal trade volume, we compiled a list of biggest trade routes[5] ranked by trade volume in both 2013 and 2014, shown in Table 1. There are 38 routes in 2013 and 37 in 2014 with at least 0.5% of global trading volume of the year, each making up about three quarters of world's total trade. Not surprisingly, the biggest routes are from the top two exporters (Indonesia and Australia) to the top four importers (China, Japan, India and Korea)—all within the Pacific region. In fact, the eight routes between these countries are the top eight routes in both 2013 and 2014, accounting for about 44% of the global trade volume in each year.

In the Atlantic region, the biggest route is Russia to UK in 2013 (1.8% of world total) and 2014 (1.9%). South Africa to India is the biggest cross-regional route in both years, with 1.7% and 2.5% world total trade respectively. Overall, among the routes listed[6] in Table 1, the Pacific market contributes 53.5%

[5]A trading route here means a unique pair of countries. The volume of each route is the total coal export from the first country to the second.

[6]Note that these are not all the routes, therefore the percentage numbers do not add up to 100%.

A: 2013					B: 2014				
From	To	MT		Region	From	To	MT		Region
World Total		1287.1			World Total		1296.4		
Australia	Japan	124.6	9.7%	Pacific	Indonesia	India	134.5	10.4%	Pacific
Indonesia	India	116.8	9.1%	Pacific	Australia	Japan	119.7	9.2%	Pacific
Indonesia	China	89.8	7.0%	Pacific	Australia	China	93.4	7.2%	Pacific
Australia	China	88.1	6.8%	Pacific	Australia	Rep. of Korea	54.9	4.2%	Pacific
Australia	Rep. of Korea	49.9	3.9%	Pacific	Indonesia	China	49.8	3.8%	Pacific
Indonesia	Japan	37.8	2.9%	Pacific	Australia	India	46.9	3.6%	Pacific
Indonesia	Rep. of Korea	36.1	2.8%	Pacific	Indonesia	Japan	35.6	2.7%	Pacific
Australia	India	34.8	2.7%	Pacific	Indonesia	Rep. of Korea	35.6	2.7%	Pacific
Kazakhstan	Russian Fed.	25.3	2.0%	Non*	South Africa	India	31.9	2.5%	Cross
Russian Fed.	China	25.1	1.9%	Pacific**	Russian Fed.	China	25.8	2.0%	Pacific**
Russian Fed.	UK	23.4	1.8%	Atlantic	Russian Fed.	UK	24.0	1.9%	Atlantic
South Africa	India	21.3	1.7%	Cross*	Kazakhstan	Russian Fed.	22.8	1.8%	Non
Mongolia	China	18.2	1.4%	Non	Mongolia	China	19.5	1.5%	Non
Indonesia	Malaysia	17.1	1.3%	Pacific	Colombia	Netherlands	17.4	1.3%	Atlantic
Colombia	Netherlands	15.6	1.2%	Atlantic	Indonesia	Thailand	16.2	1.2%	Pacific
Russian Fed.	Rep. of Korea	14.5	1.1%	Pacific	Russian Fed.	Rep. of Korea	16.2	1.2%	Pacific
Indonesia	Philippines	14.5	1.1%	Pacific	Indonesia	Philippines	15.0	1.2%	Pacific
Indonesia	Thailand	14.3	1.1%	Pacific	Russian Fed.	Japan	14.7	1.1%	Pacific
South Africa	China	13.1	1.0%	Cross	Indonesia	Malaysia	14.5	1.1%	Pacific
Indonesia	Hong Kong, China	12.9	1.0%	Pacific	Indonesia	Hong Kong, China	12.5	1.0%	Pacific
Russian Fed.	Japan	12.5	1.0%	Pacific	USA	Netherlands	11.0	0.9%	Atlantic
USA	UK	12.3	1.0%	Atlantic	Colombia	Turkey	9.9	0.8%	Atlantic
USA	Netherlands	11.5	0.9%	Atlantic	Russian Fed.	Ukraine	9.8	0.8%	Non
Canada	China	11.1	0.9%	Cross	South Africa	Netherlands	9.4	0.7%	Atlantic
Russian Fed.	Ukraine	10.6	0.8%	Non	Canada	Japan	8.9	0.7%	Cross
Canada	Japan	10.2	0.8%	Cross	USA	Brazil	8.8	0.7%	Atlantic
Colombia	UK	9.0	0.7%	Atlantic	Russian Fed.	Turkey	8.6	0.7%	Non
Russian Fed.	Turkey	9.0	0.7%	Non	Colombia	UK	8.3	0.6%	Atlantic
Netherlands	Germany	8.3	0.6%	Non	Colombia	USA	8.3	0.6%	Atlantic
Colombia	Turkey	8.2	0.6%	Atlantic	USA	UK	8.2	0.6%	Atlantic
USA	Brazil	7.8	0.6%	Atlantic	Australia	Netherlands	7.8	0.6%	Cross
USA	Rep. of Korea	7.6	0.6%	Cross	Canada	China	7.7	0.6%	Cross
Colombia	Chile	7.6	0.6%	Atlantic	Russian Fed.	Netherlands	7.6	0.6%	Cross
Canada	Rep. of Korea	7.5	0.6%	Cross	USA	Rep. of Korea	7.5	0.6%	Cross
USA	China	7.5	0.6%	Cross	Colombia	Israel	7.2	0.6%	Atlantic
South Africa	Netherlands	7.0	0.5%	Atlantic	Canada	Rep. of Korea	6.8	0.5%	Cross
Australia	Netherlands	6.8	0.5%	Cross	Netherlands	Germany	6.6	0.5%	Non
Colombia	USA	6.7	0.5%	Atlantic					

Note(*): Cross = Cross-regional, Non = Non-seaborne.
Note(**): Part of Russia's export to China is non-seaborne. (See http://www.suekag.com /deliverydirections/)

Source: UN Comtrade database (http://comtrade.un.org/data)

Table 1: Biggest Coal Trade Routes in 2013 and 2014: the Pacific market contributes 53.5% of two year's global trade volume and top eight routes in both years while only 8.9% is from the Atlantic market

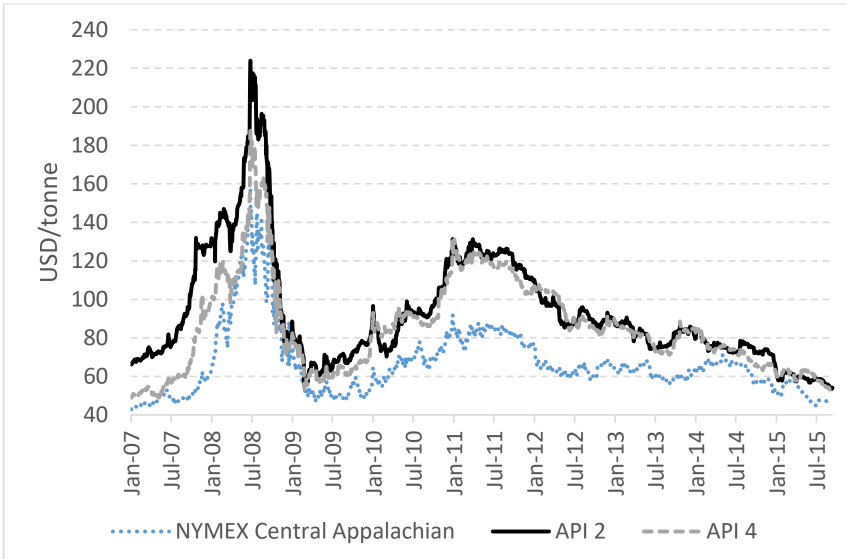

Figure 1: The three major Atlantic indexes, January 2007 to September 2015

of two year's global trade volume while 8.9% is from the Atlantic market. Cross-regional routes have a smaller presence of 5.7% of world's total trade and the rest are non-seaborne routes (such as Mongolia to China), totaling 5.4% of global volume. This illustrates the dominance of the Pacific market over the Atlantic market in physical coal trading—while the relationship is the opposite in the paper market, as said before.

As displayed in Figure 1, the NYMEX and two API indexes exhibited a spike in the middle of 2008. Figure 2, which zooms over the recent period, shows the steady decline of all three indexes since January 2014, and a rapid decline of the NYMEX index. Prices are denominated in US dollars throughout the paper as it is the case for all commodities.

3.1. Description of the data set

A summary of the coal data we use for this study is shown in Table 2. The data include prices from ten markets around the globe: five in the Pacific

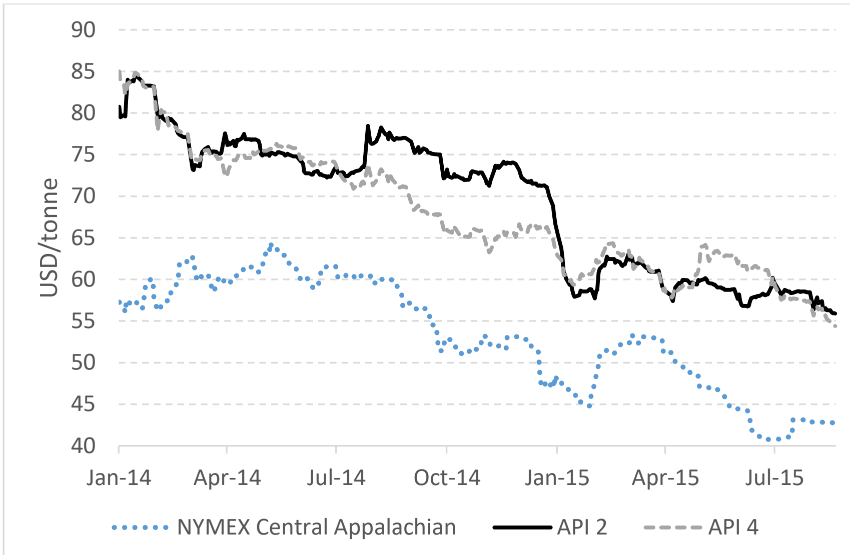

Figure 2: The three major Atlantic indexes, January 2014 to September 2015

region (Newcastle of Australia, Japan, Asia[7], China and Russia East), five in the Atlantic region (Northwest Europe ARA, Richards Bay of South Africa, Colombia, US Central Appalachian and Russia Baltic).

Different data are produced at different frequencies, from daily to monthly, and for different time ranges. ARA, Richards Bay and Newcastle have both weekly spot data for a longer time period and daily first nearby future data for a shorter period. The former are used for the PCA analysis and the latter for the cointegration analysis. All coal prices in this study are expressed in (or converted into) metric tonnes. The data is from Bloomberg.

Weekly and bi-weekly data series are converted into monthly data by averaging, as PCA requires all the variables to have the same length with no missing data. The period of analysis is the intersection of all time ranges, January 2006 to

[7]McCloskey Asian Steam Coal marker represents the CIF Asian price by equally collating prices in Korea, Taiwan and Japan.

Data source/type	Data series	Frequency	From	To
McCloskey coal marker	Japan CIF	monthly	1/31/2001	12/31/2014
McCloskey coal marker	Asia CIF	monthly	1/31/2001	12/31/2014
McCloskey coal marker	China Qinhuangdao FOB	bi-weekly	10/15/2004	1/9/2015
McCloskey coal marker	Colombia Puerto Bolivar FOB	bi-weekly	10/15/2004	1/9/2015
McCloskey coal marker	Russia West (Baltic) FOB	bi-weekly	10/15/2004	1/9/2015
McCloskey coal marker	Russia East (Vostochny) FOB	bi-weekly	10/15/2004	1/9/2015
McCloskey coal marker	NW Europe (ARA, or API2) CIF	weekly	12/16/2005	1/9/2015
McCloskey coal marker	Richards Bay (RB, or API4) FOB	weekly	10/15/2004	1/9/2015
McCloskey coal marker	Newcastle FOB	weekly	10/15/2004	1/9/2015
NYMEX first nearby	US Central Appalachian (CAPP)	daily	1/3/2006	9/18/2015
ICE first nearby	daily	7/17/2006	9/18/2015	
ICE first nearby	Richards Bay (RB, or API4)	daily	7/17/2006	9/18/2015
ICE first nearby	Newcastle	daily	1/2/2009	9/18/2015

Table 2: A summary of all coal data series for this study. The data include prices from ten markets around the globe—five in the Pacific region, five in the Atlantic region. Different prices are from different sources, available for different frequencies and time periods.

December 2014. There are 10 variables with 108 months, plotted in Figure 3. From the figure, we see similar features: a huge spike in mid-2008, another peak throughout 2011 and gradual drops afterwards.

In the cointegration analysis, we do *pairwise* cointegration tests for all pairs of prices for two main reasons: (i) we want to understand in depth the intricate dynamics between coal markets, especially by comparing intra- and inter-regional integration; (ii) we want to take advantage of the uneven nature of the data set by using the maximum amount of data available for each pair, which can be weekly or daily, and can be longer than the period used in the PCA.

Out of the ten data series we use, three are CIF (ARA, Asia and Japan) and the others are FOB[8] indexes. Wårell (2005, 2006) used only import prices; Li et al. (2010) used only export prices. We elected to follow Papież & Śmiech (2013, 2015) and use both CIF and FOB prices as the Japan and Asian ones are CIF and are too important to be left aside; the API 2 is also crucial given the gigantic amount of derivatives written on it.

[8] CIF: Cost, Insurance and Freight; FOB: Free on Board

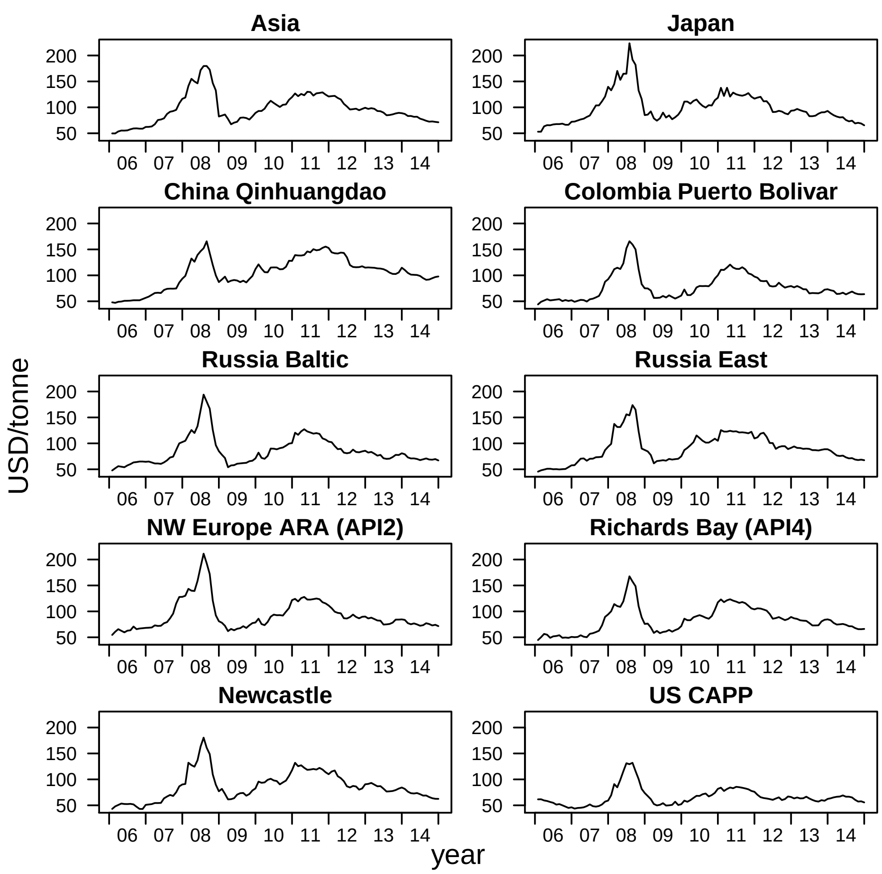

Figure 3: Price trajectories of all the ten coal data series over the period January 2006 to December 2014. All exhibited a spike in 2008.

Variable	Level	First difference
Asia	-2.4112	-7.1171 **
Japan	-2.6569 +	-5.5062 **
China	-2.3100	-6.6623 **
Colombia	-2.7715 +	-5.7417 **
Russia Baltic	-2.8570 +	-5.9577 **
Russia East	-2.3285	-7.6185 **
ARA	-1.7388	-12.5377 **
Richards Bay	-1.6591	-23.5162 **
Newcastle	-0.6877	-37.7006 **
US CAPP	-2.2368	-44.9284 **

+, * and ** denote significance at better than 10%, 5% and 1% respectively

Table 3: Augmented Dickey–Fuller test statistics

3.2. Unit root tests

As a first step of any integration study, stationarity needs to be tested. A series is called integrated with degree d, i.e., an $I(d)$ process, if it is stationary after being differenced exactly d times. $I(1)$ processes, also called unit root processes, are the most common financial time series.

We apply the augmented Dickey–Fuller (1979) test to all 10 of our variables, on both the level and the first difference. As shown in Table 3, the tests fail to reject the hypothesis that the data have unit roots at 5% level for all the variables, but reject the hypothesis that the first differences have unit roots. Therefore, all series are unit root processes as expected.

3.3. Principal component analysis

In order to see if and how the world coal market is integrated, we first perform a principal component analysis (PCA) of the ten price series.

PCA is a statistical transformation method which has been used for a long time in financial modeling, starting with the term structure of interest rates (Litterman & Scheinkman, 1991). The first component (PC1) is the linear projection of the data on the direction with the largest variance possible in n-dimensional space; the second component (PC2) is the projection on the

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	82.20	19.18	10.24	6.90	5.34	4.72	3.76	3.39	2.63	2.28
Proportion of Variance	91.7%	5.0%	1.4%	0.6%	0.4%	0.3%	0.2%	0.2%	0.1%	0.1%
Cumulative Proportion	**91.7%**	**96.7%**	98.2%	98.8%	99.2%	99.5%	99.7%	99.8%	99.9%	100.0%

Table 4: PCA results: All ten price series. The first component alone explains over 90% of the total variance. The first two components explain over 95%.

direction with the largest remaining variance, subject to the constraint that it is orthogonal to the first direction; and the process is iterated. PCA loadings are the linear transformation coefficients. For a system with n data series, there are n PCA loadings, each being a vector of length n.

A classical way to analyze PCA results is to see how many principal components are needed to explain a certain large amount, usually 95%, of total variance in the data. If most of the variance can be explained by only one variable, then the original variables are mostly integrated. Secondly, we can plot the first two or more principal component loadings on the factorial plane to see if there is grouping due to regional divide. PCA has been used in energy market integration analyses (Siliverstovs et al., 2005; Zaklan et al., 2012).

Our results in Table 4 with all 10 price series show that the first component itself explains 91.7% of the total variance; while 96.7% are explained by the first two components. This suggests that world coal markets are not fully integrated as two major factors are needed to represent the price series. In the factorial plane (Figure 4, top panel), all the variables except the US have similar PC1 loadings while their PC2 loadings spread out much more, with China being distant to others. A grouping of the eight variables is evident from the plot.

A likely reason for the separation of US prices from most others is the nature of the price series. All prices used for the PCA are McCloskey coal marker for FOB (export) or CIF (import) prices, whereas the US CAPP (Central Appalachian) is the first nearby price for an exchange-traded contract. The delivery zone of CAPP futures contract is in a tri-state area of Ohio, Kentucky

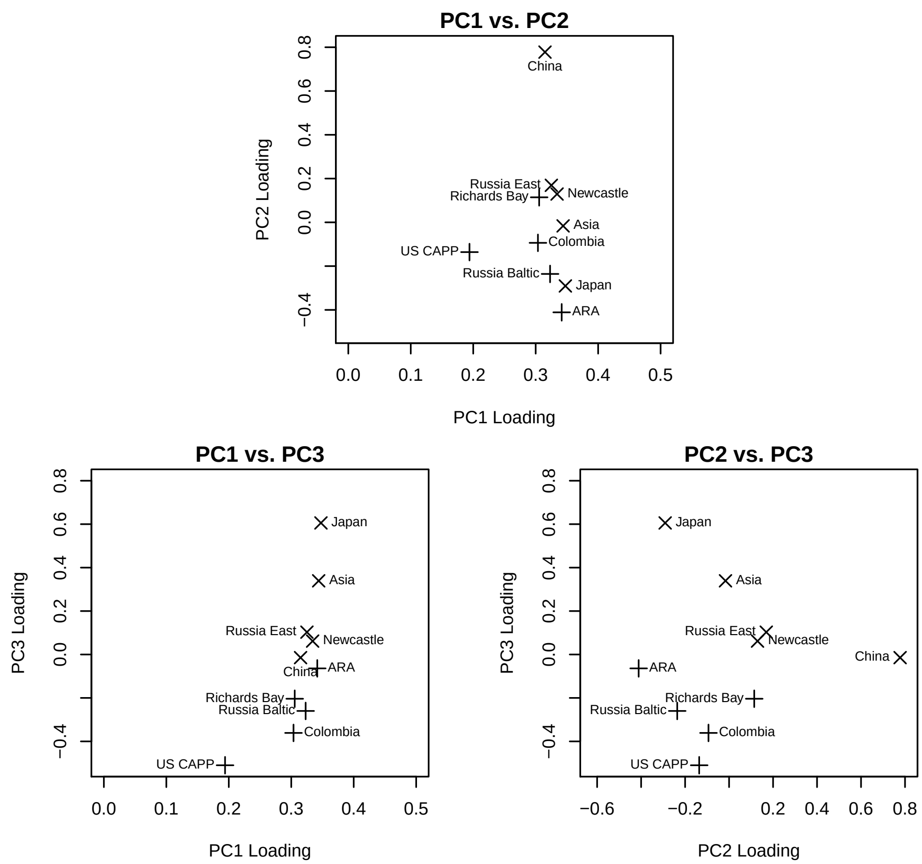

Figure 4: The first three principal component loadings of all ten price series: PC1 vs. PC2 (top panel), PC1 vs. PC3 (bottom left), and PC2 vs. PC3 (bottom right)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Standard deviation	76.53	11.82	9.27	5.40	3.99	3.61	2.80	2.32
Proportion of Variance	95.2%	2.3%	1.4%	0.5%	0.3%	0.2%	0.1%	0.1%
Cumulative Proportion	**95.2%**	97.4%	98.8%	99.3%	99.6%	99.8%	99.9%	100.0%

Table 5: PCA results: No US CAPP and China. The first component by itself explains over 95% of the total variance among the eight variables.

and West Virginia[9], not a seaport like all other prices. Interestingly, these futures contracts were delisted in December 2016, as said before.

As for China, its particular situation is due to the huge domestic demand/supply and the relatively lower importance of its international trades. Additionally, the government's measures to protect the domestic coal industry such as subsidies and tariffs also play a part in the seeming deviation of China's price from other markets. The disconnect of the Chinese price from global market prices caused by these policy-induced restrictions was pointed out by Zaklan et al. (2012).

Next, we removed the US and China series, and ran the PCA again using the other eight variables. Table 5 shows that the first component alone explains over 95% of the total variance among the eight variables. In the plot of first two component loadings (which looks similar to the top panel of Figure 4 and is not displayed), all of them still have similar PC1 loadings but diverse PC2 loadings, meaning that there is no further sub-grouping among this group according to the first two components.

Lastly, we turn to the question of whether the world coal market is still divided into Atlantic region and the Pacific region, as stated by Geman (2005) and Wårell (2005). From the top panel of Figure 4, the five markets in the Atlantic region (denoted by $+$) and the other five in the Pacific region (denoted by $\times$) are mixed—there is no separation. However, on the two bottom panels of the same figure, we can see that the two regions are clearly separated by the third PCA loading. Since the third component only explains 1.4% of total

[9] More specifically, it's on the Ohio River between Mileposts 306 and 317 or on the Big Sandy River, according to the NYMEX Rulebook.

variance compared to 5.0% by the second component from Table 4, the second conclusion we draw from the PCA tests is that the global coal market is weakly separated into the Atlantic and Pacific regions. We will return to this point in the next section.

3.4. Cointegration analysis

We now perform the Johansen cointegration tests (Johansen, 1995). Specifically, we test the cointegration relationships between each pair of data series using richer data than in the PCA, to see if the results are consistent with the PCA. The Johansen procedure has been used in the literature for integration analysis not only in the coal market (Li et al., 2010; Zaklan et al., 2012; Papież & Śmiech, 2015) but also in other commodities such as natural gas (Siliverstovs et al., 2005; Geman & Liu, 2015).

3.4.1. Model specification

A K-dimensional VAR(p) process $y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \mu_0 + \epsilon_t$ can be transformed into a vector error correction model (VECM) representation

$$\Delta y_t = \Pi(y_{t-1} - \mu_0) + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \epsilon_t \qquad (1)$$

where $\Pi := -(I_K - A_1 - \cdots - A_p)$ denotes the long-run impact, $\Gamma_i := -(A_{i+1} + \cdots + A_p)$ is the short-run impact matrix for lag $i = 1, 2, \ldots, p-1$, I_K is the K-dimensional identity matrix.

We perform the *trace test*, which has null hypothesis that the cointegrating rank is smaller or equal to r for the processes y_t. The deterministic term specification we chose is that there are intercepts (or constants) and no trends in the cointegrating relations, the so-called H_1^* model in Johansen (1995). The information criterion we used to choose the lag order in model (1) is the AIC. The test procedure and statistic are standard, and can be found in references such as Johansen (1995).

Pair	**Data frequency**	**Observations**	**Lag**	**Trace statistic**	**p-value**
ARA - Newcastle	daily	1686	4	13.9571	0.3273 o
ARA - Richards Bay	daily	2293	12	20.1617	0.0517 +
ARA - US CAPP	daily	2293	12	31.9236	0.0010 **
Asia - ARA	monthly	108	6	18.1812	0.0943 +
Asia - Russia Baltic	monthly	108	4	23.5179	0.0173 *
Asia - China	monthly	108	8	18.7493	0.0797 +
Asia - Colombia	monthly	108	3	25.4958	0.0089 **
Asia - Japan	monthly	168	9	10.0946	0.6495 o
Asia - Newcastle	monthly	108	5	23.1402	0.0196 *
Asia - Richards Bay	monthly	108	5	22.3686	0.0253 *
Asia - Russia East	monthly	108	3	19.5674	0.0622 +
Asia - US CAPP	monthly	108	3	38.0208	0.0010 **
Russia Baltic - ARA	monthly	108	2	23.4326	0.0178 *
Russia Baltic - Newcastle	monthly	108	2	16.0673	0.1715 o
Russia Baltic - Richards Bay	monthly	108	2	15.3425	0.2117 o
Russia Baltic - Russia East	biweekly	247	11	21.9540	0.0291 *
Russia Baltic - US CAPP	monthly	108	5	42.2691	0.0010 **
China - ARA	monthly	108	8	18.6152	0.0829 +
China - Russia Baltic	biweekly	247	10	17.5710	0.1132 o
China - Colombia	biweekly	247	12	14.0600	0.3187 o
China - Newcastle	monthly	108	8	20.7217	0.0433 *
China - Richards Bay	monthly	108	8	23.4109	0.0180 *
China - Russia East	biweekly	247	12	15.4460	0.2030 o
China - US CAPP	monthly	108	6	20.7625	0.0427 *
Colombia - ARA	monthly	108	2	20.9074	0.0408 *
Colombia - Russia Baltic	biweekly	247	12	24.3674	0.0131 *
Colombia - Newcastle	monthly	108	2	18.6736	0.0815 +
Colombia - Richards Bay	monthly	108	2	14.9338	0.2458 o
Colombia - Russia East	biweekly	247	12	16.9001	0.1367 o
Colombia - US CAPP	monthly	108	5	37.4549	0.0010 **
Japan - ARA	monthly	108	7	11.4845	0.5336 o
Japan - Russia Baltic	monthly	108	4	19.3481	0.0665 +
Japan - China	monthly	108	8	15.6361	0.1924 o
Japan - Colombia	monthly	108	9	24.6132	0.0120 *
Japan - Newcastle	monthly	108	7	19.0692	0.0724 +
Japan - Richards Bay	monthly	108	7	11.1782	0.5591 o
Japan - Russia East	monthly	108	8	12.6843	0.4334 o
Japan - US CAPP	monthly	108	5	34.4643	0.0010 **
Newcastle - US CAPP	daily	1686	4	17.5662	0.1134 o
Richards Bay - Newcastle	daily	1686	4	15.3299	0.2127 o
Richards Bay - US CAPP	daily	2293	12	25.4088	0.0092 **
Russia East - ARA	monthly	108	2	19.9178	0.0557 +
Russia East - Newcastle	monthly	108	6	33.7189	0.0010 **
Russia East - Richards Bay	monthly	108	2	30.6802	0.0016 **
Russia East - US CAPP	monthly	108	6	20.5340	0.0459 *

+, * and ** denote significance at better than 10%, 5% and 1% respectively; 'o' denotes that there's no cointegration between the pair

Table 6: Pairwise Johansen cointegration tests results of all the 45 pairs, showing the lag order selected for the pair, the trace test statistics and p-values for the null hypothesis $r = 0$.

		The Atlantic Region					The Pacific Region				
		Colombia	Russia Baltic	ARA	Richards Bay	US CAPP	Asia	Japan	China	Russia East	Newcastle
The Atlantic Region	Colombia										
	Russia Baltic	*									
	ARA	*	*								
	Richards Bay	o	o	+							
	US CAPP	**	**	**	**						
The Pacific Region	Asia	**	*	+	*	**					
	Japan	*	+	o	o	**	o				
	China	o	o	+	*	*	+	o			
	Russia East	o	*	+	**	*	+	o	o		
	Newcastle	+	o	o	o	o	*	+	*	**	
no. of cointegrated pairs (at 5%)		5	5	3	4	8	5	2	3	4	3

+, * and ** denote significance at better than 10%, 5% and 1% respectively; 'o' denotes that there's no cointegration between the pair

Table 7: Pairwise Johansen cointegration tests significance levels of all the 45 pairs presented in a matrix form, for easy comparisons of intra- and inter-regional pairs.

3.4.2. Cointegration tests results

In order to shed some light on the relationships among different markets, and to utilize maximal data for each pair, we start with the pairwise Johansen cointegration tests. There are 10 data series and 45 pairs. The full results are displayed in Table 6. The trace test statistics and p-values shown are for $H_0 : r = 0$. A small p-value indicates that the null hypothesis is rejected, i.e., there is a cointegration relationship between the two data series.

To have a better reading of the results in Table 6, significance levels are presented in Table 7 in a matrix form. We first note that China only has three cointegration relationships at 5%, and it is the only market with no cointegration relationship at the 1% level. This confirms the observation from the PCA test that China has a weak link with the rest of the world. The US, on the other hand, was cointegrated with all but one markets, despite the fact that it is somewhat separated in the PCA plots. One possible explanation for this is again the nature of the US CAPP index. Other reasons for the stronger link of US prices and the rest of the world may be the central location of the US and the fact that, being the only exchange-traded index, it might have been used by hedgers from both Pacific and Atlantic regions at times where the US played an important role in the coal market.

	Data series	Obs	Lag	H0: r≤0 Trace stat	p-value	H0: r≤1 Trace stat	p-value	H0: r≤2 Trace stat	p-value
Global system	10	108	18	357.637	0.0010	239.858	0.0010	156.542	0.1969
Atlantic system	5	108	6	128.557	0.0010	65.169	0.0043	24.827	0.4441
Pacific system	5	108	8	94.087	0.0018	52.031	0.0753	26.609	0.3395

Table 8: Multivariate Johansen cointegration tests results: the Atlantic region is more integrated than the Pacific region. The global coal market is still integrated between the two regions.

The bottom row in Table 7 shows the number of cointegration relationships at 5% of each market with the other nine markets. Each of the five Atlantic markets has 5 cointegration relationships on average, compared with 3.4 for the Pacific markets. If we look at the intra-regional pairs, the Atlantic region appears to be more integrated than the Pacific region. Seven of the ten intra-Atlantic (upper left quadrant of the table) and only three of the ten intra-Pacific pairs (lower right quadrant) are cointegrated at 5% level. The Atlantic market is more integrated than the Pacific despite having a much smaller trade volume, probably because of the more mature derivatives and paper trade markets. Beside the aforementioned US CAPP index, the five Atlantic prices also include the API 2 (ARA) and API 4 (Richards Bay) which are the two most important coal indexes. On the other hand, the Pacific data does not include one of its biggest exporter, Indonesia, and one of its largest importer, India. In addition, China's price does not fully reflect the market due to its policy, as remarked in the preceding section. Lastly we note from this table that, there are a significant number of cross-regional pairs (11 out of 25 in lower right quadrant) that are cointegrated at 5% level.

To further examine the regional integration, we carry out multivariate Johansen tests for all the ten series, as well as the two regions. The results are shown in Table 8. For the global system, the null hypothesis $r \leq 1$ is rejected at 5% level while $r \leq 2$ is not rejected, indicating a cointegration rank of 2. Similarly, the Atlantic and Pacific systems have cointegration ranks of 2 and 1 respectively. From this we confirm that the Atlantic region is more integrated. We also conclude that the global coal market is still integrated between the two regions, confirming the results of Zaklan et al. (2012).

By using both pairwise and multivariate cointegration tests, we are able to take advantage of the merits of both methods. For a dataset with many series (10 in our study), multivariate test results may not be reliable due to the high dimensionality, or the so-called "curse of dimensionality" (Hendry, 1995). On the other hand, pairwise tests have their own redundancy problem. It was shown by Stock & Watson (1988) that in a system with n data series and $k < n$ stochastic trends, there are $n - k$ cointegration vectors. With 10 prices and 45 total pairs in our study, at least $45 - 9 = 36$ pairs are redundant, since there are at most 9 cointegration vectors. One of the benefits of pairwise tests is that it provides us with insight into individual market dynamics, as well as inter-relationship of markets in the Atlantic and Pacific basin. Another advantage of using bivariate over multivariate tests is that it allows us to utilize the maximal amount of data for each pair. Considering all the above points, we conducted bivariate tests for all the 45 pairs first, then multivariate tests for the global system and each basin. This is similar to the two step procedure in Asche et al. (2012).

4. The Way Forward for Coal Markets

In order to check that coal price indexes' dynamics are in line with news about recent consumption numbers, we perform in this section a lead/lag investigation of six pairs of major (Eastern, Western) indexes as well as the analysis of US pure coal players' equities.

4.1. The situation of the US Giant Miners

The slow collapse of coal consumption and prices took all US publicly listed coal miners Arch Coal, Alpha Patriot Resources, Patriot Coal and Walter Energy into the protection of Chapter 11. Peabody, the largest one, followed in April 2016. The combined shares of all these companies lost a total of $30 billion in market value between 2010 and 2016. Regarding Peabody, its shares

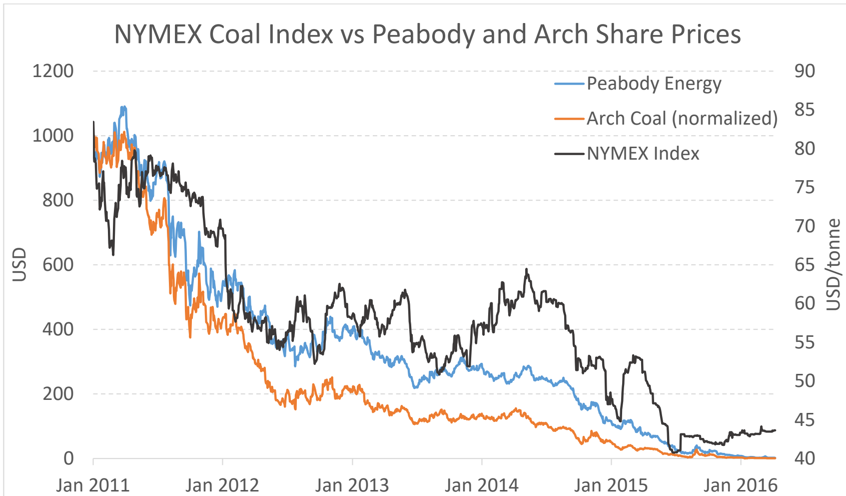

Figure 5: The share prices of Peabody Energy (split adjusted) and Arch Coal (split adjusted and normalized) versus NYMEX coal index (right axis) from January 2011 to April 2016

reached \$20 billion in 2011 and were worth \$38 million at the end of March 2016. Note that, in contrast to global miners like BHP Billiton, Rio Tinto or Anglo-American, these four companies are pure coal players.

The coal industry has been crushed by the simultaneous decline in steel production, the replacement of thermal coal by natural gas in fossil-fueled plants in a context of abundant and cheap natural gas added by the shale boom and new federal regulations on emissions.

Geman & Vergel Eleuterio (2013) exhibited in the case of fertilizer mining companies the 'leverage effect' provided by investing in their shares after the spike in wheat and corn prices during the agricultural spike in 2008, followed by a spike in fertilizer prices and generating a very large 'alpha' above the Security Market Line. We exhibit the same property in Figure 5 by plotting the decline in the major US coal index[10] (one of the few worldwide on which futures contracts

[10] As stated before, CME Group delisted CAPP futures contracts in December 2016. For hedging purposes, market participants are now using the similar but financially settled CSX

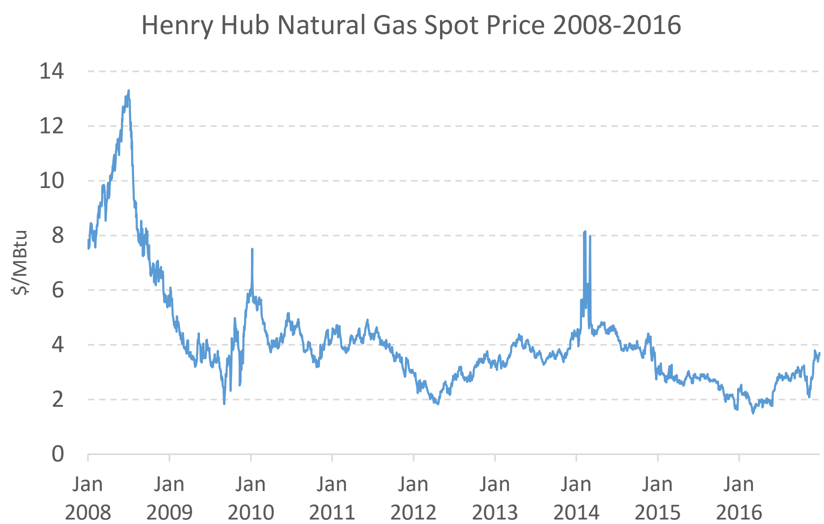

Figure 6: Henry Hub Natural Gas Spot Price 2008–2016: the trajectory depicts the low values of the Henry Hub natural gas index created by the shale gas explosion in the US after 2013, with prices mean-reverting around 3.5 dollars per Million BTUs.

are traded) and the two largest US coal companies[11].

The post- 2016 US election

Experts say that the new president's promise to save the coal industry after years of bankruptcies and bleak prospect for coal miners is not realistic. The industry is forced to compete with the declining prices of alternatives like natural gas and renewable energy sources (see Figure 6).

Federal regulations have been pushing the US away from fossil fuels, and the steps taken by a number of states to reduce the amount of emissions seem to be non-reversible. According to the Energy Information Agency (EIA), 32% of the

contracts which are based on Coal Index published by Platts. The liquidity on the CSX contracts was however still very low (an open interest of 355) as of February 2017.

[11]Both stock prices are split adjusted. Peabody had a 1 for 15 reverse spliton October 01, 2015. Arch Coal had a 1 for 10 reverse split on August 04, 2015. Arch Coal prices are also normalized by a factor of 3 to be on the same scale as Peabody prices.

Figure 7: Peabody Energy Share Price in 2016: the continuing collapse of Peabody share price during the first 10 months of 2016, the momentary rebound after the US presidential election and the return to decline in December 2016.

US electricity will come from coal in 2016, down from 50% in 2016. The shares of Peabody spiked before and just after the change of president, but were back below their value of the beginning of the year 2016 as shown in Figure 7.

The numbers are clear: according to the EIA, US mines produced about 900 million tons of coal in 2015, a 25% decline since 2008 and the lowest production since 1986; utilities and energy companies were at the same time building their natural gas and renewable capacity.

Beyond the US presidential election in 2016, experts recognize that rapidly changing factors in the energy industry more than government policy have driven the decline of coal production/consumption. The development of fracking has dramatically expanded the availability of natural gas, making this much cleaner fuel cheaper than coal in most cases.

4.2. The developments in Europe

We will illustrate the presently blurred situation in Europe with respect to coal with the examples of two countries, Germany and the Netherlands.

The year 2015 was historic for the German power sector as for the first time, renewables accounted for 32.5% of the countrys electricity production and became the number one source of power, ahead of lignite with 26%. The year-on-year increase of 20 Terawatthours was also the highest on record, with a remarkable electricity output from renewables of 31.6 Terawatthours, ensuring that the 2025 goal for the share of renewables in consumption would be reached in 2017-2018.

Coal-fuelled plants are still resisting retirement, though, because of the decommissioning of nuclear plants (such as the 1345-megawatt Grafenrheinfeld power plant in June 2015). The other reason is the record amount of electricity produced in Germany, which exported in 2015 to Austria, the Netherlands, France and Switzerland 60 Terawatthours produced in particular by coal-fuelled power plants. Political action at the level of the country and Europe will be a decisive element, to respect the Paris climate agreement of December 2015 in particular.

In the Netherlands, the Parliament voted in September 2016 in favor of a 55% cut in CO2 emissions by 2030, which would require the closure of all the country's coal-fired plants. The vote, supported both by the Liberal and Labour parties, would bring the country in line with the COP (Conference of the Parties) agreement mentioned above, with some of the most ambitious climate policies in Europe.

Five Dutch coal-fired power stations were closed in 2015, but the country still has another five plants in operation. In fact, three of them came online in 2015 – built respectively by the German and French giant utilities Eon, RWE and Engie – and have been blamed for a 5% rise in the country's emissions in 2015.

The other major event of 2016 in the Dutch energy landscape was the announcement in June 2016 by the Minister of Economic Affairs of a further reduction in the country's natural gas production. The EU largest natural gas-producing country (in particular because of the discovery in 1959 of the giant on-shore gas field of Groningen) decided to cap its annual production at 24 billion cubic meters. It was the recommendation of the State Supervision of Mines and the Dutch transmission system operator Gasunie, as a response to the significant increase of the seismic activity in the country over the years 2014 and 2015.

In a parallel manner, electricity production from coal increased in 2015 for the fourth year in a row, reaching 39 million Megawatthours and a 35% increase compared to 2014. Electricity from natural gas dropped by 9% to 46 million Megawatthours, its lowest level since 1996; electricity from renewables increased from 11.7 million Megawatthours in 2014 to 13.7 million in 2015.

4.3. The giant consumers in the East

4.3.1. China

Coal consumption in China doubled during the decade 2004 to 2014, reaching more than four billion tonnes; it then declined by 3.7% in 2015 and 2.9% in 2016 under the effect of 'hard landing' of the economy and pollution plaguing some of China's biggest cities. However, the National Energy Administration stated in November 2016 that coal consumption is projected to be stronger over the next five years than previously expected. Coal-fired electricity should go from 900 Gigawatthours in 2015 to 1100 Gigawatthours in 2020. New power capacity of all sources is increasingly sited west and transmitted to the central-Easter regions of the country. Power generation and coal chemical use in Western provinces should reach a national peak by the mid- 2020s; 'peak carbon', the peak in carbon emissions, should occur before 2030, as agreed during the Paris COP 21.

4.3.2. India

Coal continues to be a major source of energy for India, representing 44% of the primary energy mix in 2015. The power sector consumed a share of more than 60% or 527 million tonnes in 2015. Cement production, while a large consumer (5% of the total), tends to switch to alternatives such as pet coke; however, economic growth will likely mean higher cement production, as it happened in China in the early 2010s and coal requirement needs. Steel and sponge iron production do not have alternatives to their coal requirements.

India's coal reserves are the largest in the world, but the decline of the quality and the target of emission controls have created a need for more imports: 85 to 87% of India's imported coal comes from Indonesia, 10% from South Africa. Close to 20% of the population remain with limited access to electricity, and policies enacted to challenge this situation will increase India's appetite for coal. The country coal emissions are expected to soar to 5 gigatons by 2030.

4.4. Lead/Lag analysis of pairs of major (Eastern, Western) coal indexes

In this section, we perform lead-lag relationship analysis using the Information share metrics introduced by Hasbrouck (1995). Information share is a standard measure of price discovery for multiple markets. It apportions weights to individual markets, indicating the contribution of each market to the price discovery of all markets in the long run.

For this analysis, we use five daily data series, since longer sampling intervals can render the information share results unreliable (Mizrach & Neely, 2008). The five daily series are the four in Table 2 and Nymex Indonesia first nearby which has a shorter history thus not included for the PCA and cointegration study. To find out the lead-lag relationship between the Pacific and Atlantic markets, we examine the six cross-regional pairs between the three Atlantic (ARA, Richards Bay, US CAPP) and two Pacific markets (Australia, Indonesia).

We chose the period of analysis to be August 2011 to December 2015. The information share is computed daily using data of the past year (see Aggarwal, 2015). Therefore, results are reported from August 2012 to December 2015. We present the results both in Figure 8 as the dynamics in the whole period and in Table 9 as the averages in each year. From both the figures and the table, the Atlantic markets started out leading in all the pairs (except Richards

	2012	2013	2014	2015
ARA	**81.3%**	**79.3%**	**77.9%**	34.7%
Australia	18.7%	20.7%	22.1%	**65.3%**
ARA	**85.2%**	**86.4%**	**84.4%**	42.1%
Indonesia	14.8%	13.6%	15.6%	**57.9%**
Richards Bay	23.3%	**58.5%**	**52.1%**	38.2%
Australia	**76.7%**	41.5%	47.9%	**61.8%**
Richards Bay	**88.6%**	**83.4%**	**72.4%**	23.9%
Indonesia	11.4%	16.6%	27.6%	**76.1%**
CAPP	**78.2%**	**54.5%**	**78.9%**	10.6%
Australia	21.8%	45.5%	21.1%	**89.4%**
CAPP	**81.3%**	**67.0%**	**59.6%**	14.5%
Indonesia	18.7%	33.0%	40.4%	**85.5%**

Table 9: Information share results by year: the Atlantic markets led from 2012–2014; the Pacific markets overtook the lead in 2015.

Bay/Australia) and remained in the lead position in years 2013 and 2014, many of them quite dominantly. Interestingly, the situation is completely reversed in 2015, all six pairs showing a clear lead by the Pacific indexes, and a total retreat of the US CAPP index versus both Indonesia and Australia.

5. Conclusion

We revisited in this paper the integration of the world coal market through the statistical analysis of a comprehensive data set of ten indexes related to ten important countries, importers and exporters, over the past decade. The empirical tests show that China does not belong to the global market, probably because of its large production and national policies. For the US, the Principal Component Analysis shows a deviation from the other eight markets, while cointegration tests provide evidence of integration with the other eight markets. Our proposed explanation for this contradiction is the central location of the US between the Atlantic and Pacific oceans and the existence of futures contracts on the US index.

The PCA also shows that the other eight markets are integrated as a global system. The separation between the traditional Atlantic and Pacific basins is visible in the PCA only when the less important third components are displayed.

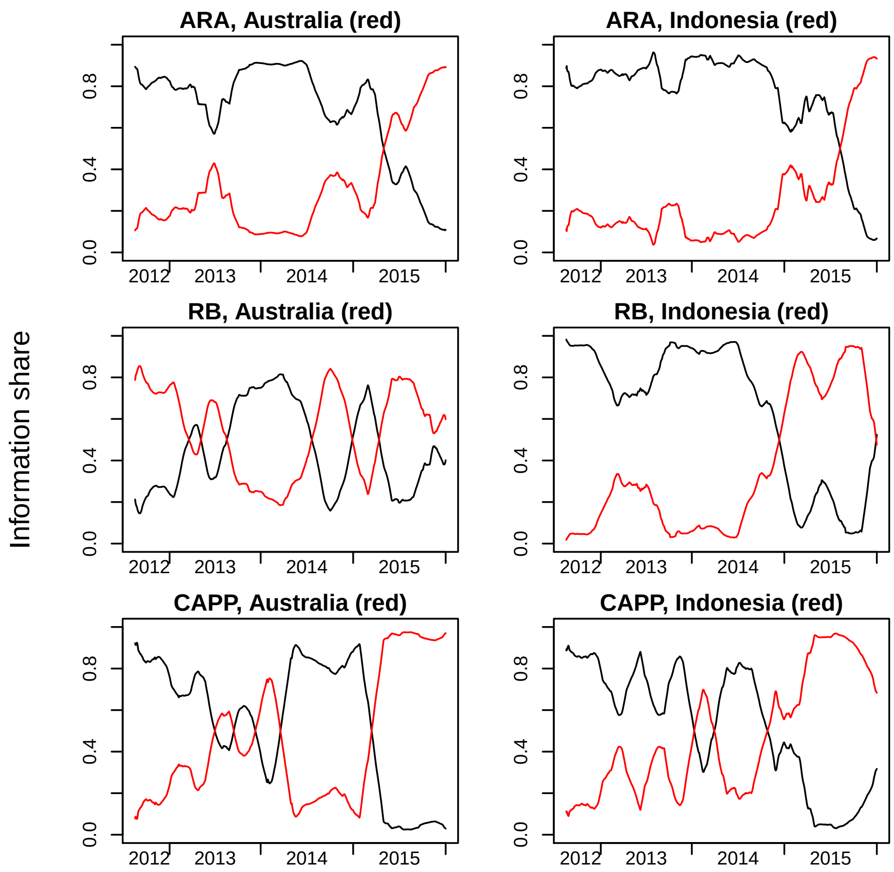

Figure 8: Information share of the six Atlantic-Pacific pairs, 2012–2015. Smoothed using three-month moving average. The Pacific market in five of six pairs started out near bottom, but all surged in 2015.

Cointegration tests lay some support to this conclusion: there are several cross-regional cointegrated pairs; the multiple tests also confirm the global system is integrated, because of cross-regional routes such as the one from South Africa to India. This confirms the results of Li et al. (2010); Zaklan et al. (2012); Papież & Śmiech (2015).

In a second part of the paper, we argue that the center of gravity of coal markets is moving East from a triple perspective: i) by reporting the remarkable reversal in 2015 of the lead/lag relationship between six pairs of major indexes; ii) by analyzing the dramatic demise of the four US giant coal miners as of 2013—the behavior of their share prices through the end of 2016 shows that this collapse was somehow due to the policy of the previous administration in terms of emissions reduction but primarily to the gigantic production of shale gas and the replacement of thermal coal by cheap natural gas in the production of electricity; iii) by describing the numbers of coal consumption across the world moving forward.

References

Aggarwal, N. (2015). *ifrogs R package*. IGIDR Finance Research Group. URL: `https://rdrr.io/rforge/ifrogs/`.

Asche, F., Gjølberg, O., & Guttormsen, A. G. (2012). Testing the central market hypothesis: A multivariate analysis of Tanzanian sorghum markets. *Agricultural Economics*, *43*, 115–123.

BP (2016). BP energy outlook to 2035, . Http://bp.com/statisticalreview.

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

Doane, M. J., & Spulber, D. F. (1994). Open access and the evolution of the US spot market for natural gas. *Journal of Law and Economics*, *37*, 477–517.

Ekawan, R., & Duchêne, M. (2006). The evolution of hard coal trade in the Atlantic market. *Energy Policy*, *34*, 1487–1498.

Ekawan, R., Duchêne, M., & Goetz, D. (2006). The evolution of hard coal trade in the Pacific market. *Energy Policy*, *34*, 1853–1866.

Geman, H. (2005). *Commodities and commodity derivatives: Modeling and pricing for agriculturals, metals and energy*. Chichester, UK: Wiley Finance.

Geman, H., & Liu, B. (2015). Are world natural gas markets moving toward integration? Evidence from the Henry Hub and National Balancing Point forward curves. *Journal of Energy Markets*, *8*, 47–65.

Geman, H., & Vergel Eleuterio, P. (2013). Investing in fertilizer-mining companies in times of food scarcity. *Resources Policy*, *38*, 470–480.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*, *50*, 1175–1199.

Hendry, D. F. (1995). *Dynamic econometrics*. Oxford University Press on Demand.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford, UK: Oxford: University Press.

Li, R., Joyeux, R., & Ripple, R. D. (2010). International steam coal market integration. *Energy Journal*, *31*, 181–202.

Litterman, R. B., & Scheinkman, J. (1991). Common factors affecting bond returns. *The Journal of Fixed Income*, *1*, 54–61.

Mizrach, B., & Neely, C. J. (2008). Information shares in the US treasury market. *Journal of Banking & Finance*, *32*, 1221–1233.

Papież, M., & Śmiech, S. (2013). Causality-in-mean and causality-in-variance within the international steam coal market. *Energy Economics*, *36*, 594–604.

Papież, M., & Śmiech, S. (2015). Dynamic steam coal market integration: Evidence from rolling cointegration analysis. *Energy Economics*, *51*, 510–520.

Sauer, D. G. (1994). Measuring economic markets for imported crude oil. *The Energy Journal*, *15*, 107–123.

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*, 603–615.

Stock, J. H., & Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association*, *83*, 1097–1107.

Wårell, L. (2005). Defining geographic coal markets using price data and shipments data. *Energy Policy*, *33*, 2216–2230.

Wårell, L. (2006). Market integration in the international coal industry: A cointegration approach. *The Energy Journal*, *27*, 99–118.

Zaklan, A., Cullmann, A., Neumann, A., & von Hirschhausen, C. (2012). The globalization of steam coal markets and the role of logistics: An empirical analysis. *Energy Economics*, *34*, 105–116.