A pair-wise analysis of the law of one price: evidence from the crude oil market *

Monica Giulietti†
Warwick Business School
University of Warwick
United Kingdom

Ana María Iregui‡
Unidad de Investigaciones
Banco de la República
Colombia

Jesús Otero§
Facultad de Economía
Universidad del Rosario
Colombia

November 2014

Abstract

This paper applies a pairwise approach to investigate the validity of the law of one price in the crude oil markets. Price differentials appear smaller between crude oil pairs with similar physical/chemical characteristics and also for pairs within OPEC.

JEL Classification: C33; F32; Q4, L7.

Keywords: Law of one price; crude oil prices; product heterogeneity; cross-section analysis.

---

*This paper was started while Jesús Otero was a Visiting Fellow in the Department of Economics at the University of Warwick. He would like to thank the Universidad del Rosario for its financial assistance (Project FIUR-195). We are also grateful to Jeremy Smith and participants at the Economic Modelling and Forecasting Seminar at Warwick Business School for useful comments and suggestions on an earlier version of the paper. The usual disclaimer applies.

†E-mail: monica.giulietti@wbs.ac.uk

‡E-mail: airegubo@banrep.gov.co

§E-mail: jesus.otero@urosario.edu.co

1 Introduction

The law of one price (LOOP) is a key ingredient in several theoretical models in economics. Because of its importance, a number of papers have assessed the validity of the law, either by testing its absolute version, according to which the prices of identical products traded in different locations must be the same once they are converted to a common currency, or by looking at its relative version, according to which the law holds when price discrepancies can be best described as stationary (or mean-reverting) processes; see, inter alia, Froot and Rogoff (1995) and Sarno and Taylor (2002).

Empirical tests of the LOOP have often investigated the prices of homogenous products traded in geographically separated markets, so that price differences can be ascribed to transportation costs (e.g. see Parsley and Wei (1996) and Goldberg and Verboven (2005)); however both the absolute and the relative version of the LOOP can be applied to diversified products if we adopt Stigler and Sherwin (1985)'s definition of a market as a product space defined by product characteristics rather than geographical distance.

This paper examines the validity of the absolute version of the LOOP in crude oil prices. The specific hypothesis that we wish to test can be stated in very simple terms, namely that the more similar the physical/chemical composition of two crude oil varieties is, the smaller their corresponding price differential. Our testing strategy also allows us to examine the role, if any, of institutional arrangements in the form of membership of the Organisation of the Petroleum Exporting Countries (hereafter OPEC). Importantly on the basis of this analysis we are able to estimate a higher bound for the extent of transaction costs operating in what has been identified by several authors as an integrated market of 'one big pool' (e.g. see Fattouh (2010)).

The paper is organised as follows. Section 2 presents a brief description of our econometric modelling strategy. Section 3 describes the data and summarises our findings. Finally, section 4 concludes.

2 Econometric modelling strategy

Our empirical modelling strategy uses the Pesaran (2007)'s pair-wise approach to testing for output and growth convergence. Adapting the notation used in that paper to our purposes, we refer to $p_{i,t}$ and $p_{j,t}$ as the spot prices of crude oil varieties i and j at time t, and define their corresponding price differential as $p_{ij,t} = p_{i,t} - p_{j,t}$, where $i = 1, ..., N-1$, $j = i+1, ..., N$, and N is the total number of crude oil varieties under consideration. The idea is then applied to investigate the order of integration of all possible $(N (N-1)/2)$ price differentials that can be constructed with a given set of N prices. For our specific purposes, we employ the ADF, ADF_{max} and ERS unit root tests developed by Dickey and Fuller (1979), Leybourne (1995) and Elliot et al. (1996), respectively, to assess the time-series properties of the resulting price differentials, as discussed in more detail in a related work (see **?**).

The pairwise approach offers the advantage that, by calculating all possible price differentials, it does not involve the choice of a benchmark variety with respect to which all other varieties ought to be measured. For the analysis of crude oil prices, this feature is indeed advantageous as recent contributions to the literature challenge the role of some historical varieties as price leaders or benchmarks. For instance, Wlazlowski et al. (2011) highlight the growing importance of the Mediterranean Russian Urals as a potential candidate benchmark variety, despite the fact that it was introduced as late as 1990. Also, Kao and Wan (2012) find support for the view that the benchmark status of West Texas Intermediate (WTI) has changed over time, to the extent that its value no longer appears to reflect underlying conditions in the market.

Once the order of integration of the oil price differentials, $p_{ij,t}$, is determined, we focus on those pairs for which the unit root null is rejected, using any of the tests listed above. For these stationary differentials, we compute the average price differential over the sample period, which we denote as

$\bar{p}_{ij}$. Notice that here we drop the subindex t since we focus on the price differentials that are stationary, and for these the mean and variance do not change over time. The average price differential is considered in absolute terms because we are only interested in its magnitude, and not in its sign.

Our hypothesis of interest is that oil price differentials, as measured by $\bar{p}_{ij}$, ought to be smaller the more homogeneous are any two oil varieties. There are two main measures for the degree of homogeneity of crude oil varieties that have been considered elsewhere in the related literature (see Wlazlowski et al. (2011)). The first one is the density of the crude oil variety i relative to variety j as measured by the absolute value of the difference between their corresponding American Petroleum Institute (API) degree content, denoted $dapi_{ij} = |api_i - api_j|$. The second measure is the absolute value of the difference between the sulphur contents of the crude oil varieties i and j, denoted $dslp_{ij} = |slp_i - slp_j|$.

In addition to the physical/chemical characteristics of different oil varieties, institutional aspects of the market, such as the existence of price agreements involving members of the OPEC cartel, may also exert some influence on price differentials, via agreed adjustments in production levels. To capture this institutional feature we introduce the dummy variable $opec_{ij}$, which is equal to 1 when varieties i and j are both produced by countries that belong to OPEC (0 otherwise), as well as the dummy variable $nopec_{ij}$, which is equal to 1 when varieties i and j are both produced by countries that do not belong to OPEC (0 otherwise).

Lastly, to allow for the possibility of non-linear effects we also included the second and third powers of $dapi_{ij}$ and $dslp_{ij}$, and the interaction between them, that is $(dapi_{ij} \times dslp_{ij})$. However, only the latter turned out to be statistically significant. The resulting cross-section regression model is:

$$\begin{aligned} |\bar{p}_{ij}| = \beta_1 + \beta_2 dapi_{ij} + \beta_3 dslp_{ij} + \beta_4 (dapi_{ij} \times dslp_{ij}) \\ + \beta_5 opec_{ij} + \beta_6 nopec_{ij} + \epsilon_{ij}, \end{aligned} \tag{1}$$

In this model, a positive sign on the coefficient associated with $dapi_{ij}$, $dslp_{ij}$ and/or $(dapi_{ij} \times dslp_{ij})$ would indicate that an increasing degree of dissimilarity of two oil varieties is associated with a wider price differential, a result which would provide support for the validity of the absolute version of the LOOP.

3 Data description and discussion of results

The database consists of the free on board (FOB) spot prices (in US dollars per barrel) of 32 crude oil varieties.[1] The price data are observed with a weekly frequency from January 3, 1997 to November 11, 2011, for a total of T =776 time observations, and are considered in logarithms. The choice of the sample period is dictated by the need to assemble the largest possible uninterrupted time series on crude oil prices. This is an updated version of the database recently employed by Wlazlowski et al. (2011) in their examination of causality relations in crude oil prices (which covered the much shorter period until March 31, 2006) and by Ghoshray and Triforova (2014), whose data set stops at January 7, 2011.[2]

The empirical analysis starts off by examining the time-series properties of the $\frac{32\times 31}{2} = 496$ crude oil price differentials. For this, we performed the unit root tests listed above including an intercept in the test regression, and

[1]The varieties are (in alphabetical order): Abu Dhabi Murban, Algeria Saharan Blend, Angola Cabinda, Asia Dubai Fateh, Australia Gippsland, Cameroon Kole, Canada Heavy Hardisty, Canadian Par, China Daqing, Colombia Caño Limon, Ecuador Oriente, Egypt Suez Blend, Europe (Ekofisk, Norway) Blend, Europe (Forcados, Nigeria), Europe (UK) Brent Blend, Indonesia Minas, Kuwait Blend, Libya Es Sider, Malaysia Tapis Blend, Mediterranean (Russia, Urals), Mediterranean Sidi Kerir Iran Heavy, Mediterranean Sidi Kerir Iran Light, Mexico Isthmus, Mexico Maya, Nigeria Bonny Light, Oman Blend, Qatar Dukhan, Saudi Arabia Heavy, Saudi Arabia Light, Saudi Arabia Medium, Venezuela Tia Juana and West Texas Intermediate (WTI) OK Cushing. Data on Canadian Heavy Hardisty is available from Jul 06/2007, so prior to this date we augment the series with Canadian Lloyd Blend which has the same API density.

[2]The data were downloaded from *www.economagic.com*. An alternative source that used to publish these data was the Energy Information Administration (EIA) of the US government at *www.eia.gov*; however, the EIA website stopped updating this information.

selecting the optimal number of lags with the Akaike information criterion (AIC), with $p_{\max} = 16$ lags. Inference is carried out at both the 5% and 10% significance levels. This follows the related work by **?**, which yields the results reproduced here in Table 1 for convenience. As can be seen in this table, at the 5% significance level both the ADF and $\mathrm{ADF}_{\max}$ tests yield rejection frequencies of 87.9 and 92.5%, respectively. Using a 10% level, the percentages are respectively 90.5 and 94.8%. For the ERS test, the percentages of rejection are 83.7 and 88.9% at the 5 and 10% significance levels, respectively. The fact that the majority of the price differentials are stationary support the view of a highly integrated oil market or, in the terms of Adelman (1984), that the crude oil market is "one great pool".

OLS estimation of equation (**??**) aims to identify the drivers of the observed price differentials across crude oil varieties by relating them to their physical/chemical characteristics and to the institutional arrangements in the market. The results are presented in Table 2. To assess the robustness of our findings, we report three regression models which vary according to the number of observations used for estimation. More specifically, in the second column of Table 2 we use information on all 496 price differentials (irrespective of the stationarity of the price differentials). Then, in the fourth column we use information on the 449 and 436 pairs which are stationary according to the ADF test at the 10% (top) and 5% (bottom) significance levels. Lastly, in the sixth and eight columns we report the results when unit-root inference is based on the $\mathrm{ADF}_{\max}$ and ERS tests.

In all regressions the coefficient of determination entails that over two-thirds of the variance in the dependent variable is accounted for by the model. The (heteroskedasticity robust) F-statistic rejects the null hypothesis that all regression coefficients (except the intercept) are equal to zero.

The estimated coefficients on $dapi_{ij}$ and $dslp_{ij}$ have the expected positive sign and are statistically significant. Thus, there is support for the hypothesis that a higher degree of similarity between oil varieties, that is relatively low values of $dapi_{ij}$ and/or $dslp_{ij}$, are associated with a smaller price differential between crude oils. In addition to this, institutional factors also appear to determine the variability in oil price differentials. Indeed, the estimated coefficient on $opec_{ij}$ is negative and statistically significant, indicating that when two crude oil varieties are produced by OPEC member countries the price differential is smaller, relative to the price differential between crude oils produced in two countries only one of which belongs to OPEC. By contrast, the estimate on $nopec_{ij}$ is positive and significant indicating that for varieties produced in non-OPEC countries price differentials tend to be larger compared to pairs of crude oils produced in two countries only of one which belongs to OPEC.

Our results indicate that nearly 70% of the variation in price differentials is accounted for by the variation in the explanatory variables, therefore only about 30% of the observed variation could be ascribed to factors others than the physical and chemical characteristics of the crude oil or institutional arrangements which affect pricing strategies.

4 Concluding remarks

In this paper we have investigated the factors which affect the size of the price differentials across 32 varieties of crude oil traded in international markets. We find that stationary crude oil price differentials can be explained to a large extent on the basis of the physical/chemical characteristics and membership of OPEC, a result which provides partial support for the absolute version of the LOOP. However, about 30% of the price differences is not accounted for on the basis of known characteristics of the crude oils and could be ascribed to transaction costs occurring within a relatively well integrated market.

References

Adelman, M. (1984). International agreements. *Energy Journal 5*, 1–9.

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

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

Fattouh, B. (2010). The dynamics of crude il price differentials. *Energy Economics 32*, 334–342.

Froot, K. and K. Rogoff (1995). Perspectives on PPP and long-run real exchange rates. In G. M. Grossman and K. Rogoff (Eds.), *Handbook of International Economics. Volume 3*, pp. 1647–1688. Amsterdam: North Holland.

Ghoshray, A. and T. Triforova (2014). Dynamic adjustment of crude oil price spreads. *Energy Journal 35*, 119–136.

Goldberg, P. and F. Verboven (2005). Market integration and convergence to the law of one price: Evidence from the European car market. *Journal of International Economics 65*, 49–73.

Kao, C.-W. and J.-Y. Wan (2012). Price discounts, inventories and the distortion of the WTI benchmark. *Energy Economics 34*, 117–124.

Leybourne, S. (1995). Testing for unit roots using forward and reverse Dickey-Fuller regressions. *Oxford Bulletin of Economics and Statistics 57*, 559–571.

Parsley, D. and S. Wei (1996). Convergence to the law of one price without trade barriers or currency fluctuations. *Quarterly Journal of Economics 108*, 1211–1236.

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

Sarno, L. and M. Taylor (2002). *The Economics of Exchange Rates*. Cambridge: Cambridge University Press.

Stigler, G. J. and R. A. Sherwin (1985). The extent of the market. *Journal of Law and Economics 28*, 555–585.

Wlazlowski, S., B. Hagströömer, and M. Giulietti (2011). Causality in crude oil prices. *Applied Economics 43*, 3337–3347.

Table 1: Proportion of stationary oil price differentials

Unit root test	α	$\bar{Z}$
ADF	5%	87.9%
ADF	10%	90.5%
$\mathrm{ADF}_{\mathrm{max}}$	5%	92.5%
$\mathrm{ADF}_{\mathrm{max}}$	10%	94.8%
ERS	5%	83.7%
ERS	10%	88.9%

Note: The underlying unit-root test regressions include a constant, and the number of lags of the dependent variable that are included in the test regression is selected using the Akaike information criterion with $p_{max} = 16$ lags. All tests are performed at significance level α. Source: **?** and authors' calculations.

Table 2: Determinants of crude oil price differentials

Variable	All pairs		Only ADF		Only ADF_{max}		Only ERS	
	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
				$\alpha = 10\%$				
Intercept	0.010	(0.004)	0.009	(0.004)	0.009	(0.004)	0.009	(0.005)
$dapi_{ij}$	0.003	(0.001)	0.003	(0.001)	0.003	(0.001)	0.003	(0.001)
$dslp_{ij}$	0.026	(0.004)	0.026	(0.004)	0.027	(0.004)	0.027	(0.004)
$dapi_{ij} \times dslp_{ij}$	0.002	(0.000)	0.002	(0.000)	0.002	(0.000)	0.002	(0.000)
$opec_{ij}$	-0.009	(0.004)	-0.009	(0.004)	-0.009	(0.004)	-0.009	(0.004)
$nopec_{ij}$	0.010	(0.004)	0.009	(0.005)	0.010	(0.005)	0.009	(0.005)
Obs.	496		449		470		441	
R-squared	0.705		0.701		0.703		0.704	
Wald F-statistic	154.858	[0.000]	134.621	[0.000]	147.402	[0.000]	143.125	[0.000]
				$\alpha = 5\%$				
Intercept			0.009	(0.005)	0.010	(0.004)	0.009	(0.005)
$dapi_{ij}$			0.003	(0.001)	0.003	(0.001)	0.003	(0.001)
$dslp_{ij}$			0.027	(0.004)	0.027	(0.004)	0.028	(0.004)
$dapi_{ij} \times dslp_{ij}$			0.002	(0.000)	0.002	(0.000)	0.002	(0.000)
$opec_{ij}$			-0.009	(0.004)	-0.009	(0.004)	-0.011	(0.004)
$nopec_{ij}$			0.009	(0.005)	0.010	(0.005)	0.010	(0.005)
Obs.			436		459		415	
R-squared			0.699		0.703		0.710	
Wald F-statistic			129.483	[0.000]	142.655	[0.000]	138.881	[0.000]

The dependent variable is measured in absolute terms. The column labelled Only ADF (ADF_{max}, ERS) indicates that only the pairs that turn out to be stationary based on the ADF (ADF_{max}, ERS) unit-root test at the 10% (top panel) and 5% (bottom panel) significance levels are used in the test regression. Standard errors are heteroskedasticity consistent. Numbers in squared brackets indicate the probability values of the robust F-test for the joint significance of all slope coefficient.