Cointegration in Crude Prices

Cameron Pfiffer

July 17, 2017

Contents

1 Introduction **2**

2 Literature Review **2**
2.1 Cointegration 2
2.2 Trading Strategy and Transaction Costs 2

3 Statistical Tests **3**
3.1 Data 3
3.2 Lag Length Selection 3
3.3 Stationarity 3
3.4 Testing for Cointegration 4

4 Trading Strategies **6**

5 Out of sample test **7**

References **7**

1 Introduction

Determining whether global oil markets are cointegrated is an important question, as doing so would allow a market participant to formulate trading strategies in cases where short-term relationships break down. It makes economic sense for the two oil grades to be cointegrated – they are essentially substitutes, with similar specific gravities and sulfur content. Indeed, Adelman (1984) claimed that "The world oil market, like the world ocean, is one great pool." This study reviews existing literature, examines the statistical relationship between Brent crude and West Texas Intermediate (WTI) crude, and proposes a trading strategy based on cointegration between the two securities.

2 Literature Review

2.1 Cointegration

Cointegration as a concept was first introduced by Granger (1981), and later expanded upon to include methodology for testing cointegration (Engle and Granger, 1987). The theory states that non-stationary series integrated of the same order may have a cointegrating vector that reduces the resulting series to a stationary process, in which case the series are said to be *cointegrated*. Assets linked by a fundamental relationship have cause to be cointegrated, and the crude oil markets are no exception. WTI and Brent crudes are both highly similar commodities, with proximate specific gravities and sulfur contents. It makes sense that two nearly fungible assets would trade at similar levels, ignoring cost of carry and transit costs. Johansen (1991) later expanded upon the Engle-Granger methodology to expand to multivariate cointegrating relationships, and offered a solution to the issue of declaring one variable dependent on the other. This paper uses both methodologies to test for cointegration.

There are consistent findings amongst academics that WTI and Brent are cointegrated, and that oil levels are $I(1)$, but stationary in first differences, consistent with expectations. Hammoudeh et al. (2008) found cointegrating relationships in Brent and WTI (as well as Dubai and Maya crudes) using both the Engle-Granger method (1987) and the M-TAR approach (Enders and Granger, 1998; Enders and Siklos, 2001) which permits asymmetry in the return to equilibrium levels[1]. Fattouh (2010) finds similar cointegrating relationships, though notes the presence of structural breaks and non-stationarity in crude price differentials in some extraordinary time periods. Numerous other academics have confirmed the existence of cointegration (Azar and Salha, 2017; Gülen, 1997; Kim et al., 2009; Reboredo, 2011) using a variety of methodologies.

2.2 Trading Strategy and Transaction Costs

Developing a trading strategy to requires the awareness of several key facts. First, in a world absent transaction costs, any deviation from the mean should spur immediate mean reversion no matter how small, as noted by Balke and Famby (1997). However, as we do not live in a world absent transaction costs and this is an empirical study (and thus subject to the whims of reality), an appropriate transaction cost must be assumed to evaluate the profitability of any trade. In this study, transaction costs are included by construction, as the trading simulation assumes that all purchases are make on the ask and all sales made on the bid. This is an overly simplified assumption, as this merely accounts for explicit transaction costs, and not implicit costs such as opportunity cost and market impact, among others[2].

Kawasaki et al. (2003) study contrarian and momentum trading strategies in cointegrated equities. Given that integrated series cross the mean frequently, the authors note that there are two general outcomes; first,

[1] Hammoudeh (2008) notes that there is cause to suspect asymmetrical returns to equilibrium; he cites the heterogeneity in global trader's expectations, compulsive and noisy trading, and transaction costs as potential factors. Asymmetry is highly important for trading strategies, though for simplicity, the cointegrating relationship between Brent and WTI is assumed to be symmetric.

[2] See Almgren and Chriss (2001) and Kissel (2006) for expanded transaction cost analysis methods.

that the spread between assets might continue away from the mean due to momentum; second, that an asset will revert towards the mean. They find Sharpe ratios of approximately 0.90 for the momentum strategy, and 0.59 for the contrarian strategy, though it should be noted that the contrarian strategy had far higher average returns, 10.69% versus the momentum strategy's 6.46%.

3 Statistical Tests

3.1 Data

The data used are daily close bid, ask, and midpoint prices for Dated Brent FOB and West Texas Intermediate FOB between January 1st, 1995 and July 14th, 2017, composing a total of 5,571 observations. Days when prices for either security was not available are removed, so that only days when both securities are traded are used. Furthermore, the data is segmented into two parts; an in-sample and out-of-sample set. The cointegrating relationships, trading strategy, and other estimates are performed on the in-sample data between January 1st, 1995 and July 14th, 2015. The remaining two years of data is used to evaluate the trading strategy using the cointegrating vectors found in the preceding period.

3.2 Lag Length Selection

It has been noted that lag length selection can have important effects on a VAR model's impulse response and variance decompositions (Braun and Mittnik, 1993), and Brooks (2014) notes that Johansen tests and Augmented Dickey-Fuller tests can be sucestible to improper lag selection. Thus, we select a lag length that creates produces the most parsimonious model, i.e., the smallest lag length. The `vars` package for R (Pfaff et al., 2013) contains a built in function for multivariate information criterion. Table 1 presents the lag lengths suggested by four information criteria tests. The Schwarz criterion suggests a lag length of 2.

Table 1: Information Criteria Lag Length

	Lag Length
AIC(n)	5
HQ(n)	5
SC(n)	5
FPE(n)	1

3.3 Stationarity

The first step in determining whether two assets are cointegrated is to find whether they are both $I(0)$. It is important to find unit roots in the levels of each potentially cointegrated series – if only one of two series is $I(1)$ and the other is $I(0)$, there can be no cointegrting vector and no inferences about the long term relationship between the two assets can be made.

Table 2 presents summary statistics for each series using lag $k = 2$, with the results clearly showing unit roots in both series. This is the expected result; few would look at a pricing chart and state that it is stationary.

Table 2: Augmented Dickey-Fuller Summary

	Augmented Dickey-Fulley Statistic	P Value
WTI	-2.006905	0.5754210
Brent	-1.683819	0.7122433

3.4 Testing for Cointegration

Two common methodologies in testing for cointegration are the Engle-Granger methodology, and the Johansen test. The Engle-Granger procedure is simple, though it suffers from some issues that cannot be solved by large sample sizes. First, hypothesis testing about the cointegrating relationship cannot be performed using the stage 1 regression, though for the purposes of this study, this is not a large issue. The goal here is to describe the relationship, not evaluate a prior expectation or challenge a thesis. The second problem is that there may be multiple cointegrating relationships due to multiple variables, though, again, this is not an issue as we are only testing two variables. Thus, the Engle-Granger method is sufficient for explaining the relaitonship between the two series.

The Johansen test is used primarily as an educational exercise, and to compare results of the cointegrating relationship. As the primary purpose of generating the cointegrating series is to forecast returns to the mean and not simply to explain movements in a series, the Engle-Granger method is perhaps less appropriate. ==========

The results for the Engle-Granger method are as follows. The first step involves estimating a model of the form:

$$\text{wti}_t = \alpha + \beta_1 \text{brent}_t + \epsilon_t$$

Next, a unit root test is run on the ϵ term to determine that it is stationary. If the null hypothesis of stationarity is rejected, the residuals are stationary, and thus WTI and Brent are cointegrated. The residuals are then used in the error correction model below:

$$\Delta\text{wti}_t = \alpha + \beta_1 \Delta\text{brent}_t + \beta_2(\hat{\epsilon}_{t-1}) + v_t$$

Where $\hat{\epsilon}_{t-1}$ is the residuals estimated in the first step. The intuition for the model is fairly simple - once the cointegrating relationship (β_1) is estimated, the coefficient of the lagged error (β_2) represents the "speed" at which past errors are corrected. It should also be noted that an intercept (α) is included in both the preceding equations, as there is little cause to exclude the intercept from the regression, and the intercept term is insignificant anyway. Thus, if $\hat{\epsilon}_{t-1}$ is large and β_2 is negative (as expected[3]), there will be (on average) a revision towards the mean in the opposite direction of the error. Table 3 displays the results of the residuals test, and Table 4 summarises the results of the error correction model.

Table 3: Engle-Granger Residuals Test

	Augmented Dickey-Fulley Statistic	P Value
ϵ	-4.634569	0.01

Table 4: Engle-Granger Error Correction Model Summary

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept (α)	0.0001362	0.0114117	0.0119335	0.9904792
Brent (β_1)	0.8262598	0.0089063	92.7723709	0.0000000
Residuals (ϵ)	-0.0143456	0.0023672	-6.0600599	0.0000000

As Table 4 demonstrates, the coefficients on the Brent is significant, as would be expected of a cointegrated relationship, and the coefficient on the residual is negative and significant.

[3] It's highly unlikely that β_2 would ever be positive, as that represents an explosive process, and the process wouldn't be stationary.

It's also important to note that, because the error correction model is estimated using OLS, ensuring that the assumptions of the estimator are not violated is worth a cursory check.

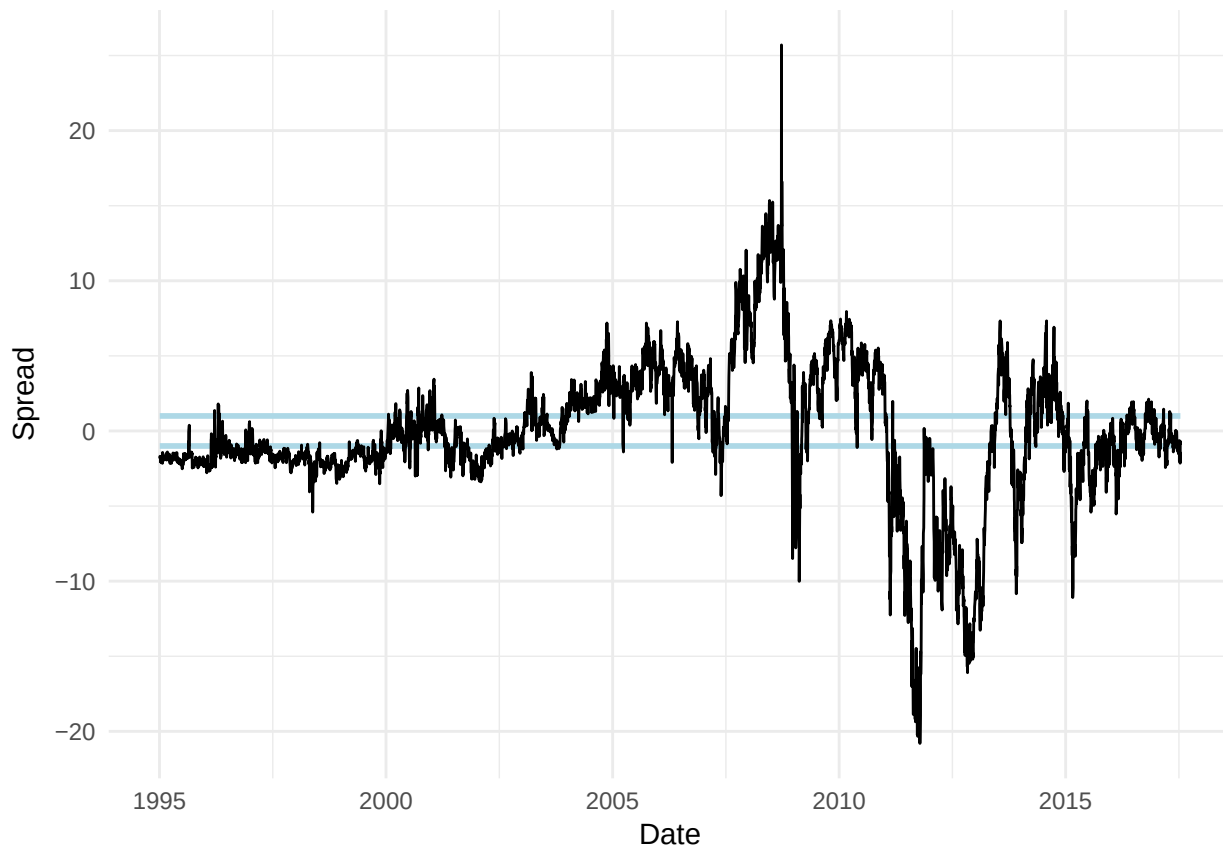

Figure 1: Cointegrated Series of WTI and Brent

4 Trading Strategies

The overview of the trading strategy is that it follows several simple steps.

1. Calculate a stationary series using a given cointegrating vector.
2. Take an appropriate long/short position when the spread is outside a threshold.
3. Close out position when the spread reaches the opposite threshold.
4. Repeat until end of series, and then close out positions.

Using the Johansen test's cointegrating vector, a spread variable is created by calculating the below equation:

$$\text{spread}_t = \alpha + \text{wti}_t - \beta_1 \text{brent}_t$$

Using the cointegrating vector found previously:

$$\text{spread}_t = -5.425 + \text{wti}_t - 0.873\text{brent}_t$$

Where spread_t is the stationary series created by an intercept (α), the price of WTI at time t, and the price of Brent at time t multiplied by a cointegrating coefficient (β). Figure 1 plots this series for the entire data set. A threshold is then utilized as in Dunis et al. (2006), whereby WTI is purchased and Brent is sold if the spread is below the negative threshold (long the spread) or the opposite if the spread is above the positive threshold.

Using this series,

5 Out of sample test

References

Adelman, M.A., 1984. International oil agreements. The Energy Journal 5, 1–9.

Almgren, R., Chriss, N., 2001. Optimal execution of portfolio transactions. Journal of Risk 3, 5–40.

Azar, S.A., Salha, A., 2017. The bias in the long run relation between the prices of brent and wti crude oils. International Journal of Energy Economics and Policy 7.

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration. International Economic Review 38, 627–645.

Braun, P., Mittnik, S., 1993. Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions. Journal of Econometrics 59, 319–341. doi:10.1016/0304-4076(93)90029-5

Brooks, C., 2014. Introductory econometrics for finance. Cambridge university press.

Dunis, C., Laws, J., Evans, B., 2006. Trading futures spreads: An application of correlation and threshold filters. Applied Financial Economics 16, 903–914. doi:10.1080/09603100500426432

Enders, W., Granger, C.W.J., 1998. Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. Journal of Business & Economic Statistics 16, 304–311. doi:10.1080/07350015.1998.10524769

Enders, W., Siklos, P., 2001. Cointegration and threshold adjustment. Journal of Business & Economic Statistics 19, 166–76.

Engle, R.F., Granger, C.W., 1987. Co-integration and error correction: Representation, estimation, and testing. Econometrica: journal of the Econometric Society 251–276.

Fattouh, B., 2010. The dynamics of crude oil price differentials. Energy Economics 32, 334–342. doi:http://dx.doi.org/10.1016/j.eneco.2009.06.007

Granger, C.W., 1981. Some properties of time series data and their use in econometric model specification. Journal of econometrics 16, 121–130.

Gülen, S.G., 1997. Regionalization in the world crude oil market. The Energy Journal 18, 109–126.

Hammoudeh, S.M., Ewing, B.T., Thompson, M.A., 2008. Threshold cointegration analysis of crude oil benchmarks. The Energy Journal 29, 79–95.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica 59, 1551–1580.

Kawasaki, Y., Tachiki, S., Udaka, H., Hirano, T., 2003. A characterization of long-short trading strategies based on cointegration, in: Computational Intelligence for Financial Engineering, 2003. Proceedings. 2003 Ieee International Conference on. IEEE, pp. 411–416.

Kim, J., Oh, S., Heo, E., 2009. A study on integration of the world crude oil markets using price asymmetry model. Geosystem Engineering 12, 1–4.

Kissell, R., 2006. The expanded implementation shortfall: Understanding transaction cost components. The Journal of Trading 1, 6–16.

Pfaff, B., Stigler, M., Pfaff, M.B., 2013. Package "vars".

Reboredo, J.C., 2011. How do crude oil prices co-move? Energy Economics 33, 948–955. doi:http://dx.doi.org/10.1016/j.eneco