Skin in the game: Resource proximity and price impact*

Tom Steffen†

January 14, 2019

Abstract

We devise a novel dataset by integrating over-the-counter oil forward trading with exchange-traded futures activity to investigate the intricate interactions between the two markets. We answer a longstanding open question and report evidence that, on an intraday basis, the futures market is the dominant information leader, but that the forward market impounds a non-negligible 20% of price innovations. Forwards are also less noisy. The futures leadership is in line with the theory and findings of Figuerola-Ferretti and Gonzalo (2010). Moreover, we use the forward market centrality of traders with substantial 'skin in the game' in the oil market as a proxy for fundamental supply and demand information. Forward trades by more central participants have a more significant price impact on the futures market of up to 15 bps over a 10-minute window.

JEL classification: G13, G23, L14, Q02, Q41.

Keywords: forwards, futures, network analysis, OTC markets, physical oil

*The Capital Markets Cooperative Research Centre (CMCRC) funded this research. We thank Platts Singapore and the Securities Industry Research Centre of Asia-Pacific (SIRCA) for facilitating the acquisition of the data. We are grateful to Riccardo De Blasis, Petko Kalev, Ognjen Kovacevic, Ian Marsh, Eugenio Piazza, Marcel Prokopczuk, and Florian Schroeder for helpful comments and discussions. We thank Nidhi Aggarwal for her public contributions to the computation of price discovery metrics in **R**, and acknowledge the use of Marek Hlavac's 'stargazer' **R** package for formatting LaTeX tables, and Francois Briatte's 'ggnet2' package for network visualizations.The views expressed in the paper are those of the authors. All remaining errors are our own.

†Corresponding author: Geneva Finance Research Institute, University of Geneva. Bd du Pont d'Arve 42, 1211 Geneva 4, Switzerland. Email: tom.steffen@unige.ch.

1 Introduction

The financialization of commodity markets, often defined as financial investors driving prices via speculation, has been a contentious topic over the last decade. In the oil market, derivatives markets (paper oil) have evolved rapidly alongside the physical markets (spot oil or cash oil), and the two are inextricably linked. In the North Sea, the different physical and financial oil contracts are commonly known as the Brent Complex.[1] This study focuses on the physically settled forward Brent contracts (also called forward BFOE, an acronym for the North Sea Brent-Forties-Oseberg-Ekofisk oil fields) and the financially settled ICE Brent Crude futures contracts to answer open questions in the literature on the importance of both contracts for the determination of the efficient price of oil. Data constraints, such as the reliance on low-frequency physical oil price proxies by previous studies (see for example Kaufmann and Ullman, 2009; Liu et al., 2015), impair our understanding of the intraday price discovery process in the oil market to this day. Our proprietary dataset allows us to analyze the intricate intraday over-the-counter (OTC) trading of physical oil, and the impact the latter has on oil derivatives.

The OTC trading of physical oil has high barriers to entry, requiring participants to receive and deliver crude oil. Hence, the market mainly attracts oil majors, such as BP, Chevron, ConocoPhillips, Shell, and Total, and commodity traders, for example Glencore, Mercuria, Phibro, Trafigura, and Vitol (see Barret, 2012, for an extended list of market participants). In a Bloomberg article by Cheong et al. (2017), commodity trading companies argue that superior information is required to trade successfully in the oil market:

> "The most valuable commodity out there is information, and the most useful information is the proprietary, critical information that you obtain from your own supply chain. You have to have skin in the game. You have to have access to assets, whether it's infrastructure, terminals, vessels or refineries."

Accordingly, not only are the major oil corporations heavily invested in the oil supply chain but commodity trading houses continuously increase their investments in infrastructure too. We hypothesize that the 'skin in the game' argument reflects the structure of the oil market, where participants in the physical market are informed, and their trading behavior impacts the futures market. Participants are intensively involved in physical trading for many reasons, but their activity is arguably often based on supply and demand fundamentals received from their upstream (exploration and production) or downstream (refining, processing, and distribution) business lines. Their trading activity therefore reasonably serves as a proxy for fundamental information. For example, many physical trading participants are also owners or operators of oil fields that are feeding into the major North Sea oil grades (so-called equity owners), run refineries, own vessels, or invest in pipelines.

[1] Elements of the Brent Complex include physical crude oil cargoes and forward contracts, Contracts for Differences (CFDs, which are short-term swaps between elements of the complex), Exchange of Futures for Physicals (EFPs, which price the differential between futures and forwards), and many others.

The idea that physical OTC trading reflects fundamental information and thus serves as a signal to futures markets aligns with the literature, as commercial companies may capitalize on their superior knowledge of physical market conditions to exploit informational frictions (Cheng and Xiong, 2014; Frino et al., 2016).

This paper is structured into three components. In a first step, we establish a cointegrating relationship between the forward and the futures markets, and then decompose their price series into permanent innovations and transitory effects in order to determine price leadership. In a second step, we take a closer look at the trading process in the forward market and why it is important for the oil price development. In a third step, we link the trading activity in the forward market, and the 'skin in the game' of its participants, to the revelation of fundamental information and its incorporation into the futures price.

With regards to the cointegration and information leadership between the physical and financial oil markets, we find that the futures market is the information leader and incorporates approximately 81% of innovations to the efficient oil price. Actually, it is the case that, during most of the day, the futures market is responsible for 100% of the price discovery. This is explained by the fact that the forward market is only active during a short period at the end of the trading day—from 16:25 to 16:30. Our study, however, demonstrates that, as soon as the forward market becomes active, even during this short period of the day, it manages to claim a non-negligible 19% of the price discovery share in the oil market. These findings align with Figuerola-Ferretti and Gonzalo (2010), who explain the permanent-transitory decomposition between the spot and financial markets, and whose results also establish futures price leadership for non-ferrous metals.

The forward BFOE market is characterized by a core-periphery structure, with a selected few core traders dominating the trading activity. Participants in the periphery interact with each other occasionally but trade more intensely with the core participants, who appear to adopt the unofficial role of 'market makers'.[2] Addressing the 'skin in the game' argument above, we hypothesize that the trading activity of core forward BFOE participants conveys information to the financial oil market and therefore significantly impacts the Brent futures price. In accordance with this proposition, a more central forward trader, as determined by the weighted out-degree network centrality measure, has a more significant price impact on the futures market—up to 15 basis points (bps) over a 10-minute window. This reaction very likely corresponds to the impounding of fundamental information from the physical crude oil market, given that the dominant traders in the forward market have infrastructure stakes, investments, and connections to the upstream and downstream crude oil supply chains.

We contribute to the literature by identifying the price discovery roles of both the futures and forward markets on an intraday level. We demonstrate that the futures market

[2]In the context of this paper, we do not use the term 'market maker' in its traditional sense of an equity stock exchange liquidity provider. We use the term in the strict sense of the Platts methodology documents, where it refers to a trading participant in their system who provides a quote before a certain cut-off period. Please refer to the institutional details in Section 2.

is unsurprisingly the information leader, but the physical market still plays an essential role in determining oil price developments. Second, we provide first-hand evidence on trading activity in the forward market, and on how the major participants in this market influence financial oil prices as well. We add to the debate on the financialization of oil and the information transmission between spot and futures, by showing that the proximity to the natural resource and oil infrastructure appears to provide physical market participants with fundamental information that is revealed via forward trading and subsequently incorporated into futures prices.

The financialization debate in the academic literature discusses how financial investors affect and potentially distort trading in commodity markets. The futures market performs two crucial roles: (i) risk sharing—commodity producers hedge their price risk in the futures market for which speculators provide liquidity; and (ii) information discovery—centralized futures trading supplements the decentralized spot trading in information discovery (Cheng and Xiong, 2014).

We focus on the price discovery role played by financially settled oil and physically settled oil. The intersection of the exchange-traded (ET) and OTC market structures of oil has been the subject of active debate for years (see Garbade and Silber, 1983). In the commodity literature, centralized futures trading is seen to facilitate information aggregation, in the sense of Grossman and Stiglitz (1980) and Hellwig (1980), by solving informational frictions arising from the complicated supply, demand, and inventory dynamics of the spot market (Cheng and Xiong, 2014). However, Sockin and Xiong (2015) argue with their model that noise in commodity futures trading can create confusion whether speculation or economic fundamentals are driving prices. Empirical evidence on price discovery is inconsistent (see Bekiros and Diks, 2008; Inci and Seyhun, 2017; Kaufmann and Ullman, 2009; Liu et al., 2015; Quan, 1992; Schwarz and Szakmary, 1994; Silvapulle and Moosa, 1999), with some reporting a unidirectional relationship from futures to spot or vice versa, others a bidirectional relationship. Most of the studies using higher-frequency data (daily), but suggest that futures prices lead the price discovery and influence the spot prices (see for example Figuerola-Ferretti and Gonzalo, 2010). The findings are not surprising given the superior futures liquidity due to contracts that are ET, financially settled, consist of smaller lot sizes, have lower transaction costs, and are not constrained by operational requirements to handle physical oil. However, most, if not all, of these studies focus on low-frequency data (daily or monthly) and use proxies (such as benchmarks) to account for the physical market, since OTC data on spot oil trading is difficult to obtain. The low-frequency characteristic is a significant shortcoming given that adjustments to shocks in these markets occur within minutes (see Inci and Seyhun, 2017).

In addition, numerous studies provide theoretical and empirical support for the assertion that commodity market financialization substantially impacts oil information discovery and price developments (see for example Basak and Pavlova, 2016; Büyükşahin and Robe, 2014; Cifarelli and Paladino, 2010; Henderson et al., 2015; Silvennoinen and Thorp, 2013; Singleton, 2013; Tang and Xiong, 2012). For instance, prices are driven by the large financial inflows into commodity futures from index investors, changes in hedge fund positions, or increased volatility and correlation with other financial indexes. Other studies endorse fundamental supply and demand as the driver of price developments (see for example Büyükşahin and Harris, 2011; Fattouh et al., 2013; Hamilton, 2009; Hamilton and Wu, 2015; Irwin and Sanders, 2011; Irwin et al., 2009; Juvenal and Petrella, 2015; Kilian, 2009; Kilian and Murphy, 2014; Knittel and Pindyck, 2016). They often reject the 'bubble claim' that prices are driven purely by speculation. Overall, Cheng and Xiong (2014) conclude that the financialization has altered commodity markets considerably.

Our study investigates the financialization of oil from the information discovery perspective. Our dataset consists of the order book of OTC forward oil contracts traded on the Platts *eWindow* platform—the most popular and active market for physical North Sea crude oil. We integrate OTC forward order book data with ICE Brent Crude futures data from Thomson Reuters Tick History (TRTH) on an intraday frequency to analyze price discovery and test our 'skin in the game' hypothesis. We thereby try to distill the effect of fundamental physical oil market information on the futures market.

The remainder of this paper is organized as follows: the next section (2) describes the institutional background. Section 3 introduces the data and provides descriptive statistics of the forward and futures markets. Section 4 presents the primary results on the price discovery of both oil contracts, trading networks, and the impact of forward transactions on the futures price. Section 5 concludes.

2 Institutional details

2.1 Platts' *eWindow*

Platts, the leading provider of reference prices in the energy markets, operates a system called the Editorial Window (eWindow) to assess the Dated Brent benchmark. The eWindow resembles an OTC trading venue consisting of a real-time open order book that reveals bids, offers, and ensuing trades. It is where price discovery takes place in the physical oil market. Frino et al. (2017) provide a more detailed description of the mechanism.

As described by Barret (2012), the final 30 minutes of Platts' so-called Market on Close (MOC) process, from 16:00 to 16:30, concentrate liquidity in the physical oil market. During the daily half-hour period, known as the Platts Window, Platts computes the Dated Brent benchmark price based on the combination of trading activity in three OTC products: (i) physical North Sea cargoes, (ii) short-term swaps between Dated Brent and Forward Brent (i.e., CFDs), and (iii) outright forward Brent (also called cash BFOE).

The interest in this study only lies in the last element, the cash BFOE contract, since it is used to trade long-term supply and demand and is the physical counterpart of the futures contract. Cash BFOE is, therefore, the most appropriate contract to focus on in our 'skin in the game' context. Moreover, because we study North Sea crude oil dynamics, we do not incorporate products from other markets into our analysis. Naturally, many factors, products, and markets globally contribute to oil price discovery but are outside the scope of this paper. In addition, Davis (2012) determines that the Platts Dated Brent benchmark prices approximately 67% of the global physical oil traded and one might argue that the trading activity of North Sea physical and financial oil reflects most of the information.

Trading in eWindow is organized and governed by Platts' rules. As such, one can either trade as a so-called 'market maker' or 'market taker'. To become a market maker during the half-hour Platts Window, a participant must indicate his interest to trade to Platts ahead of a cut-off period by submitting a new bid/offer. After the cut-off period, Platts accepts no new bids/offers, and only existing quotes can be amended.[3] However, so-called market takers can hit the bid or lift the offer of a market maker at any time. The cut-off time for cash BFOE is 16:25:00 and after that only existing quotes can be amended by the market makers. Bids/offers for the forwards can be changed until the close at 16:30:00. This five-minute phase is judged to be of critical importance for price discovery in the physical oil market. After 16:30:00 all bids/offers that have not been acted upon during the Platts Window expire.[4]

While only a limited number of companies, mastering the operational requirements of trading physical oil, participate in trading via eWindow, a more substantial number of subscribers to Platts' fee-based Global Alert (PGA) real-time information service can follow the live physical trading activity and order-flow information (transactions, bids, asks). This is of importance to this paper, since it allows, for example, futures traders to gain insights into physical oil price developments.

It is important to note that physical oil trading can take place throughout the day as well. However, the MOC methodology has the advantage of promoting liquidity in an illiquid market, as it leads to a natural concentration of activity in a short period at the end of the day (Barret, 2012). Typically, the vast majority of the daily forward quoting and trading activity is concentrated between 16:25:00 and 16:30:00 (quote amendments and trading) and some of it between 16:20:00 and 16:24:59 (quote submissions before the cut-off). Given that forwards are the physical counterparts of futures, which, however, trade throughout the day, we focus on the last five minutes of the window.

2.2 The forward market

The forward contract derives its specification from Dated Brent, commonly considered the spot price for a cargo of North Sea oil. Since January 2012, Dated Brent has reflected the price of a crude oil cargo with an assigned shipping date between 10 and 25 days ahead. Forward Brent contracts, in contrast, specify the month of loading but have no date yet assigned. The seller communicates the date to the buyer within 25 days of the delivery, and thus the contract is also called *25-day forward*. It follows that forward contract expiry

[3]Source: `http://www.rusneftekhim.com/docs/crude_oil.pdf`.

[4]Information received during the *Platts Oil Methodology Explained* session at the *Platts London Oil & Energy Forum*.

is on day number five in a 30-day calendar month (with slight deviations for longer or shorter months); for example, the May12 contract expired on 5 April 2012. After that, Jun12 would have been the active contract.[5] In February 2015, Platts extended the spot Dated Brent date range to 10-30 days ahead. This change means that the forward contract now expires on the last business day of the month following the month-ahead Dated Brent date range. For example, the May15 contract expired on 31 March 2015.[6]

Forward price changes need to be incremental (under normal market conditions from 1 ¢/barrel (bbl) to 3 ¢/bbl) and prices (denominated in USD [$]) must stand firm long enough to be acted upon by a counterparty, to ensure orderly price discovery.[7] Forward contracts can be traded up to three months ahead and are settled physically (Barret, 2012). The minimum trade size for forward BFOE is a partial cargo of 100,000 bbl. The majority of quotes correspond to this size. Occasionally quotes contain a quantity of 200,000 bbl, and can go up to 600,000 bbl (corresponding to a full cargo). The minimum shipment size acts as barrier-to-entry to the market. Only a limited number of companies, mastering the operational and logistical requirements of trading physical oil, participate in trading via eWindow. The firms are also required to satisfy Platts' due diligence requirements.

2.3 The futures market

ICE Brent Crude futures are traded on ICE Futures Europe (IFEU) and are listed for each month seven years forward. We sample only the front-month, closest-to-maturity futures contract and roll over to the next contract at expiry.[8] Futures and forward expiries did not align precisely before March 2016. This had to do with the assessment of the Dated Brent and the implications for the forward contract maturities, as explained in the previous section.

All Brent futures contract months up to and including February 2016 expired at the end of the business day preceding the 15$^{\text{th}}$ calendar day before the start of the next contract month. For example, the Feb16 contract expired on 14 January 2016. Starting with the March 2016 contract, Brent futures have expired on the final business day two months

[5] Until 5 April, the 10-25 spot date range falls within April; the forward contract is thus May. After 5 April, the 10-25 spot date falls within May, and the forward contract is thus June. See `https://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/25Day_Brent_Calendar.pdf`.

[6] Until 31 March, the 10-30 spot date range falls within April; the forward contract is thus May. After 31 March, the 10-30 spot date falls within May, and the forward contract is thus June. See `https://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/faq-month-ahead-dated-brent.pdf` and `https://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/Dated-Brent-Month-Ahead-Calendar.pdf`.

[7] Source: `https://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/oil-timing-increment-guidelines.pdf` and `https://www.platts.com/IM.Platts.Content/MethodologyReferences/MethodologySpecs/Platts-Forward-Curve-Oil.pdf`.

[8] Using only the nearest-maturity contracts is consistent with the literature on commodity derivatives. This is mainly because the closest futures contract is typically the most liquid, whereas the longer-dated contracts are predominantly thinly traded.

ahead of the contract month in question. Thus, the Mar16 contract expired on 29 January 2016.

Before March 2016, we match the front-month forward contract with the closest futures maturity at that time. For example, the Aug15 forward contract would be matched to the nearby Jul15 futures contract from 1 June 2015 to 15 June 2015 and then the nearby Aug15 futures contract until 30 June 2015. Since the March 2016 adjustment, the futures and forward expiries have aligned.

The contract size in the futures market is 1,000 bbl and thus considerably smaller than the contract size in the forward market. The currency denomination is USD ($) per bbl, and the minimum price increment is 1 ¢/bbl. The Brent futures are cash settled against the ICE Brent Index, which is computed based on forward market activity. Moreover, a close link to the physical market exists via the EFP contract which converts a Brent futures position into a physically deliverable forward contract. For these reasons, futures and forward prices commonly converge at expiry.

3 Data

Full order book data on physical oil trading was acquired from S&P Global Platts. The data consist of message-by-message activity for Platts Cash BFOE partial cargoes, also known as BFOE forward contracts. The dataset includes multiple forward maturities/contract months. We determine and focus on the front-month contract and use the data to reconstruct the full order book from 3 January 2012 to 1 February 2017, which includes trading of the contract months Feb12 to Apr17. Message timestamps are in milliseconds and the time zone is Greenwich Mean Time (GMT). We aggregate the data at the second frequency and convert all timestamps to reflect London local time.[9]

All standard order book variables, such as time, price, and quantity, are recorded and messages are labeled with a unique identifier and a sequence number, allowing us to trace the order life cycle from inception to the final state. Importantly, the forward data also contain the trader identifiers. As such, the identity of the sender of each message is known. Moreover, for transactions, the buyer and seller are reported too. Finally, the directionality of a transaction, i.e., the passive side as well as the active side of the trade, can be determined.

At the same time, Brent futures data for the same date range are obtained from TRTH. The data also include all standard variables, including the last trade price, bid and ask prices, and volumes. We sample the futures data at the second interval with timestamps reflecting London local time. The futures data do not contain participant identifiers.

We clean and merge the datasets together to create one aggregated time series of both forward and futures prices, allowing us to track the developments in both markets. Given

[9]We account for British Summer Time (BST), starting on the last Sunday of March and ending on the last Sunday of October.

the particularities of the forward market, as described in the institutional details section, there are five minutes each day during which the forward market activity overlaps with that of the futures market. To account for the registration of interest mechanism of the Platts Window (with new submissions cut-off ahead of 16:25), the window of interest extends from 16:22 to 16:30.[10]

Table 1: Summary statistics

Quotes			**Trades**	
Time	*Observations*	%	*Observations*	%
16:25-16:30	76,166	91	4,553	100
16:22-16:25	5,616	7	-	-
before 16:22	445	1	-	-
after 16:30	1,470	2	3	0
Quantity	*Observations*	%	*Observations*	%
100 K bbl	83,658	100	4,556	100
200 K bbl	34	0	-	-
300 K bbl	1	0	-	-
400 K bbl	2	0	-	-
600 K bbl	2	0	-	-
Participants				
	Total	*per maturity*	*per day & maturity*	
Forwards	22	10.46	3.70	
Transactions				
	Trading days	*Total*	*per maturity*	*per day & maturity*
Forwards	1,070	4,556	72.32	4.26
Futures	1,319	3,627,935	57,586.27	2,750.52

Note: This table reports the summary statistics for front-month forward and futures trading. For forward quotes, *Observations* count the messages recorded on the Platts platform including new quote submissions, changes, cancellations, and executions for each of the specified time windows as well as the contract sizes ranging from 100,000 bbl to 600,000 bbl. For forward trades, *Observations* count the number of executed transactions only for the same categories. *Total*, *per maturity*, and *per day and maturity* report the average number of forward participants over the full sample period, each contract month, and each trading day in a traded contract month respectively. *Trading days* reports the number of active trading days in both contracts, while *Total*, *per maturity*, and *per day and maturity* contrast the number of forward and futures transactions in our sample.

Although the Brent futures and forward markets are closely interlinked, their structures are quite distinct. For this reason, we provide some comparative descriptive statistics of the data at our disposal in Table 1. We focus on the front-month contracts.

First of all, 91% of the quoting activity in the forward market falls within the five minutes from 16:25 to 16:30. 7% falls within the period from 16:22 up until 16:25. The

[10]See Appendix A.1 for full details on the data-merging process. Moreover, there are days when Platts performs an early assessment and therefore the window of interest ranges from 12:22 to 12:30.

remaining activity occurs either before or after this. Nearly all quoted prices have a quantity of 100,000 bbl attached. Regarding trades, 100% execute for the minimum trade size of 100,000 bbl.

The requirements that must be fulfilled in order to trade in the forward market are, by nature, more restrictive than those for the futures market. Hence, the total number of participants in the forward market over our entire period of investigation amounts to 22.[11] Although we do not have participant information for the futures market, it is reasonable to assume that the number is far more significant. The average number of forward traders during each contract month is 10.46. On a daily basis, on average, only 3.70 traders participate in the front-month contract. The quoting activity of the five most active traders accounts for 53% of all quote submissions, while they make up 68% of the total number of executed transactions.[12]

From 2012 to 2017, forwards traded on 1,070 days, while futures traded on 1,319 days. A total of 4,556 front-month forwards were traded, virtually all of which were traded between 16:25 to 16:30. This corresponds to 4.26 trades per day. Overall, each contract month traded 72.32 times on average. In the futures market, during the same period and five-minute window, a total of more than 3.6 million transactions were concluded, with a mean volume of 2.06, accounting for a transaction size of roughly 2,060 bbl (for parsimony this result is not tabulated here). This is significantly less than the 100,000 bbl transaction size in the forward market. On a daily basis, this corresponds to an average of 2,750.52 front-month futures transactions, or 57,586.27 per contract month.

4 Empirical analysis

4.1 Price discovery: Does the forward market matter?

The methodology in this section is based on Baillie et al. (2002), Gonzalo and Granger (1995), Harris et al. (2002), Hasbrouck (1995), Lehmann (2002), Putniņš (2013), and Yan and Zivot (2010).

Following the notation and presentation in Baillie et al. (2002), two price series that are cointegrated $I(1)$ are denoted $Y_t = (y_{1t}, y_{2t})'$ with an error correction term $z_t = \beta' Y_t = y_{1t} - y_{2t}$, and have a cointegrating vector $\beta = (1, -1)'$.

The information share (IS) and component share (CS) are both based on a vector error correction model (VECM) of the form

$$\Delta Y_t = \alpha \beta' Y_{t-1} + \sum_{j=1}^{k} A_j \Delta Y_{t-j} + e_t \tag{1}$$

[11] This corresponds to the number of participants quoting in the market and differs from the 21 traders that completed transactions as reported later in this study.

[12] These results are not tabulated due to the need to guarantee the anonymity of the traders, consistent with the data provision license.

where the error correction vector is α; the zero-mean and serially uncorrelated innovations are termed e_t, with Ω being their covariance matrix. The first right-hand-side element in Equation 1 expresses the long-term relationship, also called the equilibrium dynamics, and the second right-hand-side element represents the short-term relationship between the two price series, driven by noise (bid-ask bounces, inventory calibrations etc.).

$$\Omega = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \tag{2}$$

Accordingly, σ_1^2 is the variance of e_{1t} and σ_2^2 of e_{2t}. ρ is the correlation between the innovations.

From Hasbrouck (1995), one can convert Equation 1 into the integrated vector moving average (VMA), as represented in Equation 3:

$$Y_t = \Psi(1)\sum_{s=1}^{t} e_s + \Psi^*(L)e_t \tag{3}$$

$\Psi^*(L)$ is a matrix polynomial with a lag operator, L. $\Psi(1)$, called the impact matrix, depicts the sum of the moving average coefficients, i.e., the cumulative impact of an innovation e_t on the price. Again, the first right-hand-side element represents the long-term price impact of an innovation, and the second expression is the transitory component, which does not have a permanent price impact. Due to the long-term impact having the same effect on both price series, the impact matrix has identical rows, denoted $\psi = (\psi_1, \psi_2)$ in the next equation:

$$Y_t = \iota\psi\Big(\sum_{s=1}^{t} e_s\Big) + \Psi^*(L)e_t \tag{4}$$

where ι is a column vector consisting of ones.

Hasbrouck (1995) shows that ψe_t is the common efficient price of the two series, also called the common factor component, impounded into prices due to information. There is a close link between Equation 4 and the Stock and Watson (1988) common trend:

$$Y_t = f_t + G_t \tag{5}$$

where the common factor component is denoted f_t and G_t is the transitory component.

Hasbrouck (1995) demonstrates that the information share of a market is the contribution of that market to the total variance of the efficient price innovations, $var(\psi e_t) = \psi\Omega\psi'$. The computation for the Hasbrouck (1995) IS, identifying market i's contribution to price discovery, is therefore

$$IS_i = \frac{([\psi M]_i)^2}{\psi\Omega\psi'},\ i = 1, 2. \tag{6}$$

where M is a lower triangular matrix. Ω is only diagonal if price innovations across markets are uncorrelated. Because Ω is often not diagonal, the Cholesky factorization of $\Omega = MM'$

is used to deal with the significant correlation of the innovations, e_t, by attributing the covariance term to the first market, leading to an upper bound estimate of the IS_i.

$$M = \begin{pmatrix} m_{11} & 0 \\ m_{12} & m_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2(1-\rho^2)^{1/2} \end{pmatrix} \tag{7}$$

The common approach is, therefore, to change the order of the price series and repeat the process, and then take the average of the lower and upper bounds to determine the IS_i (see Baillie et al., 2002). Baillie et al. (2002) show that, the higher is the correlation, the greater is the divergence between the upper and lower bound estimates. The lower bound thereby represents only the price's contribution, while the upper bound also includes the contribution from the correlation with the second price.

Equation 5 leads to the CS estimation proposed by Booth et al. (1999), Chu et al. (1999), and Harris et al. (2002) based on the Gonzalo and Granger (1995) permanent-transitory decomposition. The latter show that $f_t = \Gamma Y_t$. Γ is the common factor coefficient and Baillie et al. (2002) demonstrate that it is the orthogonal to the error correction coefficients $\alpha'_\perp = (\gamma_1,\ \gamma_2)'$.

The CS for market i can thus be computed as

$$CS_i = \gamma_i = \frac{\alpha_{\perp,i}}{\alpha_{\perp,1} + \alpha_{\perp,2}},\ i = 1, 2. \tag{8}$$

or

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1},\ C2_2 = \gamma_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2} \tag{9}$$

Equation 9 shows that, if $\alpha_i = 0$, all price discovery takes place in market i, as that market does not correct for a disequilibrium between the two price series (Yan and Zivot, 2010).

Lastly, we follow Yan and Zivot (2010) and Putniņš (2013) and calculate the information leadership share (ILS):

$$IL_1 = \left|\frac{IS_1}{IS_2}\frac{CS_2}{CS_1}\right|,\ IL_2 = \left|\frac{IS_2}{IS_1}\frac{CS_1}{CS_2}\right| \tag{10}$$

and

$$ILS_1 = \frac{IL_1}{IL_1 + IL_2},\ ILS_2 = \frac{IL_2}{IL_1 + IL_2} \tag{11}$$

The *ILS* reported in this paper is the average of ILS_1 and ILS_2. We use the *ILS* for our main inference, as Putniņš (2013) demonstrates that *IS* and *CS* diverge if the levels of noise in the two markets differ. Both metrics then measure a combination of price leadership and relative avoidance of noise. The *ILS*, however, provides a clean measure of price discovery leadership, as it cancels out the dependence on the noise component. We follow the definition in Putniņš (2013) and determine that a market is the information leader if its price is the first to reflect innovations in the fundamental value of the underlying.

We aggregate our data at the one-second frequency and do so to reduce the noise in the estimation of the price discovery measures. A higher sampling frequency leads the lower and upper bound estimations to be very close to each other. The contemporaneous correlation is negligible because the *IS* estimation can more accurately identify the sequence of the markets' responses to new information (see for example Hasbrouck, 1995, 2003; Tse, 2000).[13]

In our analysis, we determine the price discovery measures on a daily basis for each front-month contract (63 months from February 2012 to April 2017), and average across days and then months.[14] We only include days on which the futures and forward markets are cointegrated at the 75% confidence level or higher. We use the Akaike information criterion (AIC) test to determine the optimal number of lags.[15] The reason for selecting this more lenient confidence level is the paucity of forward quoting activity and therefore the difficulty in establishing cointegration at the usual levels. Based on this, in our sample, 466 trading days are cointegrated. The results are reported in Table 2 and show that price discovery takes place in both the futures and the forward markets.

Table 2: Price discovery measures

Statistic	ISFUT	ISFOW	CSFUT	CSFOW	ILSFUT	ILSFOW
Mean	0.66	0.34	0.48	0.52	0.81	0.19
Median	0.67	0.33	0.48	0.52	0.83	0.17
Min	0.19	0.12	0.19	0.35	0.41	0.01
Max	0.88	0.81	0.65	0.81	0.99	0.59
St. Dev.	0.13	0.13	0.10	0.10	0.11	0.11

Note: This table reports the mean, median, min, max, and standard deviation of the futures information share, *ISFUT*, forward information share, *ISFOW*, futures component share, *CSFUT*, forward component share, *CSFOW*, futures information leadership share, *ILSFUT*, and forward information leadership share, *ILSFOW*, respectively. The reported values are computed on a daily basis using log prices and then averaged across days and months.

The average daily information share of the futures market ($ISFUT$) across contract months amounts to 66%, while the forward market ($ISFOW$) makes up the remaining 34%. This split is not surprising given that, proportionally, much fewer quotes and transactions take place in the forward market. Generally speaking, forwards are only active for five

[13]We choose one-second intervals to minimize the computational power required to compute the price discovery measures. However, our conclusions remain unchanged if we use millisecond data.

[14]This averaging approach does not materially affect the reported means of the price discovery metrics. We do this to report meaningful minimum and maximum values by contract month. Due to the volatile nature of the price discovery estimations, daily minimum and maximum values would equal 0.01 and 0.99.

[15]We use the Trace cointegration rank test and obtain the critical values from Johansen (1995). This approach is not uncommon. For example, Figuerola-Ferretti and Gonzalo (2010) use the 80% confidence level to establish cointegration between copper futures and spot. Our results are not materially affected by choosing a higher or even lower cut-off.

minutes a day. These five minutes coincide, however, with arguably the most crucial period of the trading day in the oil market. This is when the price assessment of the Platts Dated Brent benchmark is in full swing and the spot, as well as financial, oil market is unusually alert (see for instance Frino et al., 2017).

Across contract months, the average daily component share shows a more even split between the two markets, indicating even that the forward market is leading, with the *CSFUT* accounting for 48% and the *CSFOW* for 52% of the price discovery. The results for *IS* and *CS* can differ substantially because the price series are affected by different noise levels. "*CS* values low noise relative to speed, *IS* values speed relative to low noise, and *ILS* values only speed" (Putniņš, 2013, p. 81).

The measure of interest is, therefore, the *ILS*, which cancels out the noise of the price series, as developed by Yan and Zivot (2010) and Putniņš (2013). The futures market dominates price discovery, accounting for an *ILSFUT* of 81%. Nonetheless, the *ILSFOW* still amounts to 19%, suggesting that the physical oil trading introduces innovations to the oil market on a regular basis. This finding indicates that the forward market might be slower in incorporating information but is much less noisy, leading to the 50-50 split between *CSFUT* and *CSFOW*. The result aligns with the fact that the forward-to-futures quote ratio is infinitesimal, as only a select few companies can participate in forward trading. These companies often have a direct interest in the physical oil market and close links to supply and demand fundamentals through their upstream and downstream business lines. Their activity is thus often motivated by commercial needs. The futures market, in contrast, with its many participants with diverse trading interests, is much noisier. For instance, financial investors regularly engage in speculation on future oil price movements without possessing superior information, in line with the theory on the financialization of commodity markets. However, after accounting for the differences in noise, the *ILS* confirms the *IS* result, suggesting that the futures market is the leader in reflecting innovations about the fundamental value of oil.

Table 3: Price discovery leadership

Leadership	n
Forward	61
Futures	405

Note: This table reports the information leadership on a daily basis for all front-month contracts as measured by *ILS*. n indicates the number of information leadership days of the forward and futures contract respectively.

The three average daily price discovery measures by month are volatile, as indicated by standard deviations from 10% to 13%, as well as minimum and maximum *ILS* values that vary from just 1% to 59% in the case of the forward contract. Looking at this on a day-by-day basis, the futures contract is the uncontested information leader, guiding the forward contract on 405 out of 466 days (Table 3). Figure 1 further illustrates the consistent price leadership of the futures contract over time. Based on the five-day moving average, the *ILS* of the futures contract hovers between 60% and 100%, thereby claiming the majority of the price leadership. Nonetheless, the forward contract manages to claim more than 50% of the information leadership occasionally, even though its share also regularly drops down to 0%.

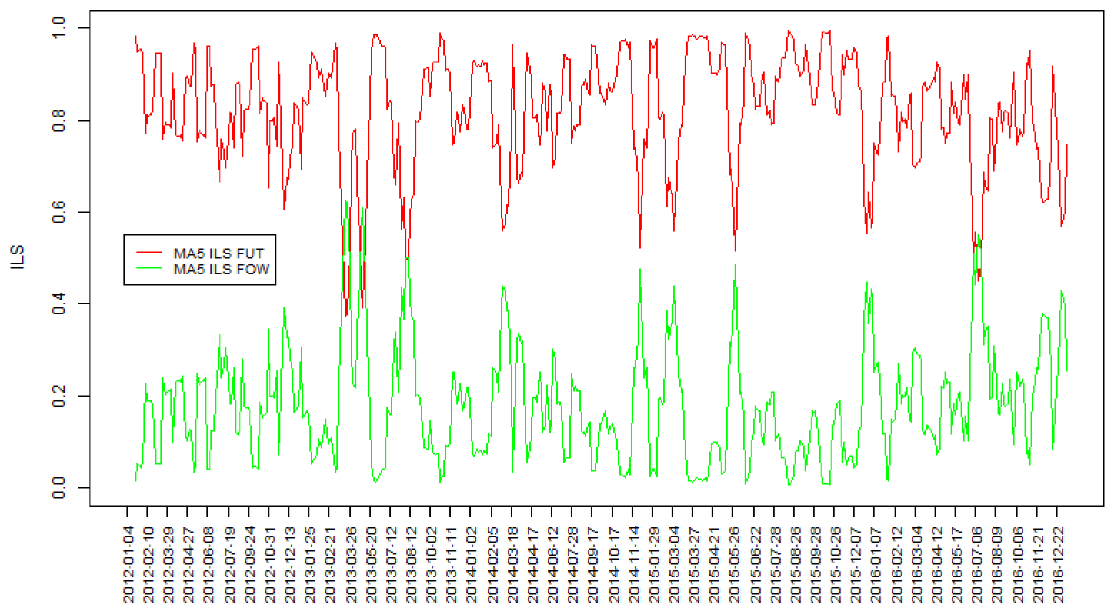

Figure 1: Price discovery over time

Note: The y-axis depicts the *ILS* ranging from 0% to 100%. The x-axis shows the date range. The red line represents the five-day moving average of *ILSFUT*. The green line represents the five-day moving average of *ILSFOW*.

While the futures contract commonly leads the forward contract, informationally, the results demonstrate that the physical and financial oil markets closely interact with each other, and both contribute to the price discovery process on a daily basis. Interestingly, however, the forward price is less noisy and reflects nearly 20% of price innovations. The futures' informational dominance is likely driven by liquidity advantages because they are exchange-traded, financially settled, trade in smaller lot sizes, and have lower operational requirements and barriers to entry.

4.2 Networks in the physical oil market

Since many in-depth academic studies look at the oil futures market (see for example Liu et al., 2015) but acknowledge that, due to data constraints, little can be said about its physical counterpart, in this section we are the first to analyze OTC forward trading more closely.[16] The obtained data allow us to address the limitations of previous studies by applying techniques from social network analysis (SNA) that have recently found their way into financial economics, tackling questions such as how networks impact returns, price discovery, information diffusion, and OTC trading (see for example Di Maggio et al., 2017a,b; Hendershott et al., 2017; Li and Schürhoff, 2014; Munyan and Watugala, 2017; Ozsoylev et al., 2014).[17]

Figures 2 and 3 depict trading in the forward BFOE market. A node (circle) represents a trader, while the edge (arrow, line) that connects two traders represents an interaction (trade). The network figures are produced with the so-called Fruchterman-Reingold force-directed layout algorithm, which determines the optimal position of nodes by simulating attractive and repulsive forces to find an equilibrium state that minimizes the energy of the system.

Traders are assigned random numbers and are labeled *Ti*. For our period, there are 21 traders (which is different from the 22 quoting participants) in the cash BFOE market, and thus $i = 1, ..., 21$. These are mainly oil majors, commodity traders, and oil explorers, operators, and refiners, but the occasional financial institution is also represented. Additionally, many of these companies are so-called equity owners in North Sea oil grades, defined as owners or operators of oil fields that feed into one of the four BFOE oil grades. This fact speaks directly to our 'skin in the game' hypothesis, as some forward traders have direct infrastructure stakes in the underlying North Sea oil market.

The node size represents the centrality of the traders in the network and is determined by the weighted out-degree measure. The measure computes the number of outgoing edges of a node, counting interactions (including multiple interactions) with other nodes. Outgoing means that the arrow illustrates the directionality, i.e., the trade flow from the passive market maker's perspective. This is important because we want the centrality measure to reflect the relevance of the party that is revealing its intentions to either buy or sell. The edge weight thus determines the strength of the relationship, meaning the number of trades initiated by one trader and acted upon by the other trader. The weighted number of outgoing edges, therefore, represents the importance of a market maker in Platts' eWindow by also taking into account its market share. Without the instigation of a market maker, no trade will take place. The centrality score, also depicted next to the figures, will be used as input to the regressions in the next section in the form of the *CENT* variable.[18]

[16]Several studies, such as those by Barret (2012) and Fattouh (2011), conduct qualitative research on the interrelations between physical and financial oil, but no quantitative analysis has been undertaken.

[17]For detailed surveys on the application of social networks in economic research, please refer to Easley and Kleinberg (2010), Goyal (2005), and Jackson (2005, 2008).

[18]The network and centrality are determined based on all forward transactions in all contract months

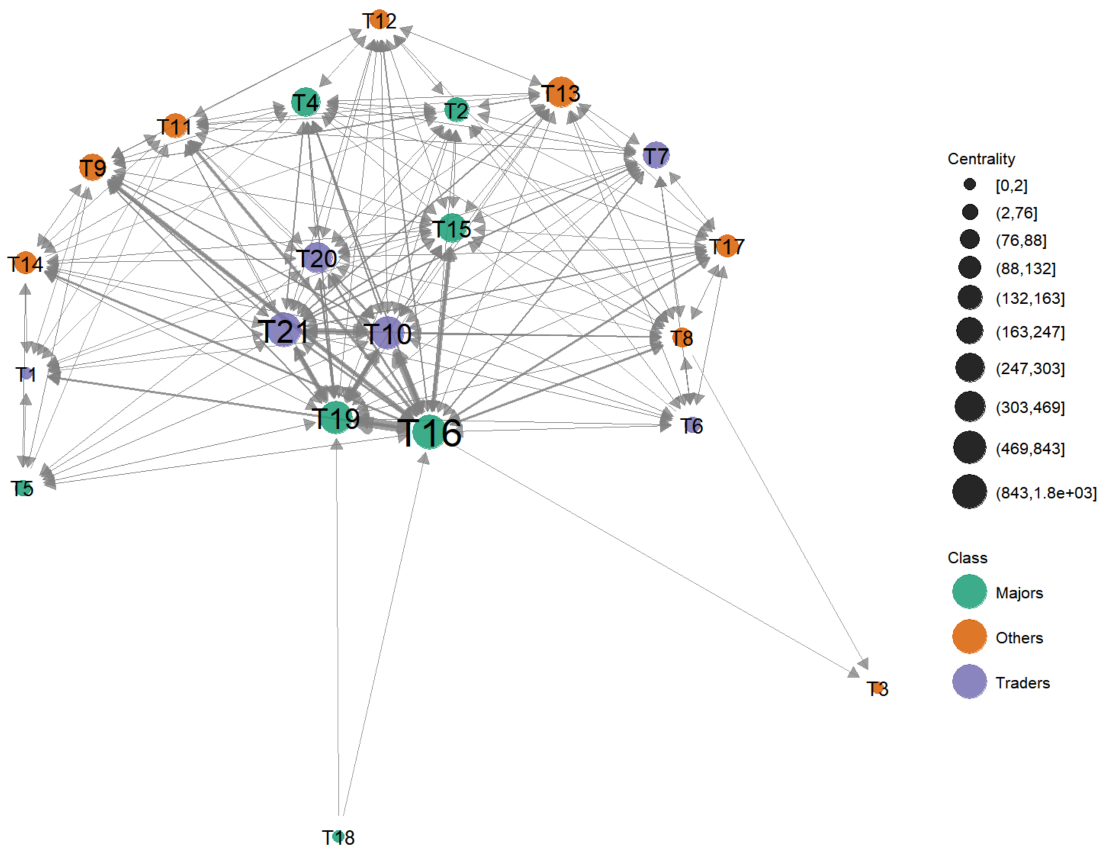

Figure 2: Forward BFOE trading network

Note: This figure depicts the trading network in all forward BFOE contract months from February 2012 to April 2017 using the Fruchterman-Reingold algorithm. Arrow directionality is determined from the view of the passive buy/sell side of the trade—the so-called market maker according to Platts' terminology. A gray outgoing arrow therefore indicates trader i passively buying from or selling to another trader, or both. Edge weights outline the strength of the relationship. The node size and its respective text size indicate the centrality of the trader as measured by the weighted out-degree, i.e., the number of outgoing edges representing the importance of the trader as a market maker. The colors for *Majors*, *Others*, and *Traders* represent the classification into oil majors, commodity trading houses, and other business lines.

over the full sample period. The reason we use the entire sample period is that we aim to measure the importance of a market maker and its reputation as a major trading participant, established over time. We use transactions in all contract months to capture the overall standing of a trader in the market. In robustness tests we use (i) a compounded yearly centrality measure and (ii) only front-month forward trades instead, and find that the centrality ranking is remarkably persistent over time and that the results remain unchanged.

Based on the weighted out-degree measure, *T16* is the most central trader, followed by *T21*, *T10*, and *T19*.

We surmise that the revealing of trading intentions by the main participants in the forward market impacts the prices in the futures market because it divulges information on the supply and demand of the actual physical resource. Although driven by a different intuition, the 'NYSE specialists literature' shows that trades with specialist participation have a higher immediate impact (see for example Hasbrouck and Sofianos, 1993). On the one hand, the futures market's reaction could stem from a mechanical relationship driven by the same participants trading in both the forward and futures markets and potentially triggering herding by other futures participants. On the other hand, trading strategies of futures traders observing physical market activity via Platts' PGA service (see Section 2.1) could drive the price impact in the futures market. In both cases, the forward market serves as a signal to the futures market.

In Figure 2, the nodes are classified into oil majors, commodity trading houses, and other auxiliary businesses such as explorers, refiners, and financial companies. The core of the trading network is dominated by oil majors (green) and commodity traders (purple), while the periphery is made up of all three categories, but mainly auxiliary companies (orange). Within the core, oil majors have strong interactions amongst each other, as can be seen by the thick arrows between *T16* and *T19* and *T15* and *T16*. However, commodity traders occupy a central role in the market, being strongly connected with each other (*T21* with *T10*), but also with the oil majors in their network vicinity (*T21* with *T19* and *T16*, and *T10* with *T19* and *T16*). Moreover, a triangular relationship can be identified between *T10*, *T19*, and *T16*. Both majors and traders within the core have many trading interactions with less central participants too.

In Figure 3, the core-periphery relationship structure of the network is highlighted. The green nodes (*T16*, *T21*, *T10*, and *T19*) build the core, and the rest of the traders are more or less peripheral. An edge adopts the color of the node if the interaction is between nodes of the same group (core-core or periphery-periphery interactions); an edge adopts the grey color for connections between nodes of different groups (core-periphery interactions). There are two 'outliers' that rarely interact with the market; trader *T3* that only has incoming edges, which means it only trades aggressively, and trader *T18* whose outgoing edges indicate its passive role in the market.

The figure underlines strong core-core trading relationships, as depicted by the thick green lines, indicating that core participants interact with each other frequently. Core-core interactions account for the majority of the trading activity. Periphery-periphery interactions are mostly weak. The thin orange arrows suggest intermittent trading in the outer perimeter of the network, indicating occasional rather than established trading relationships. There are some moderate core-periphery relationships, as illustrated by the medium-strength gray arrows between orange and green nodes. These connections imply that some peripheral participants regularly trade with the same core participants. Examples include the edges between *T15* and *T16*, *T9* and *T16*, and *T2* and *T21*.

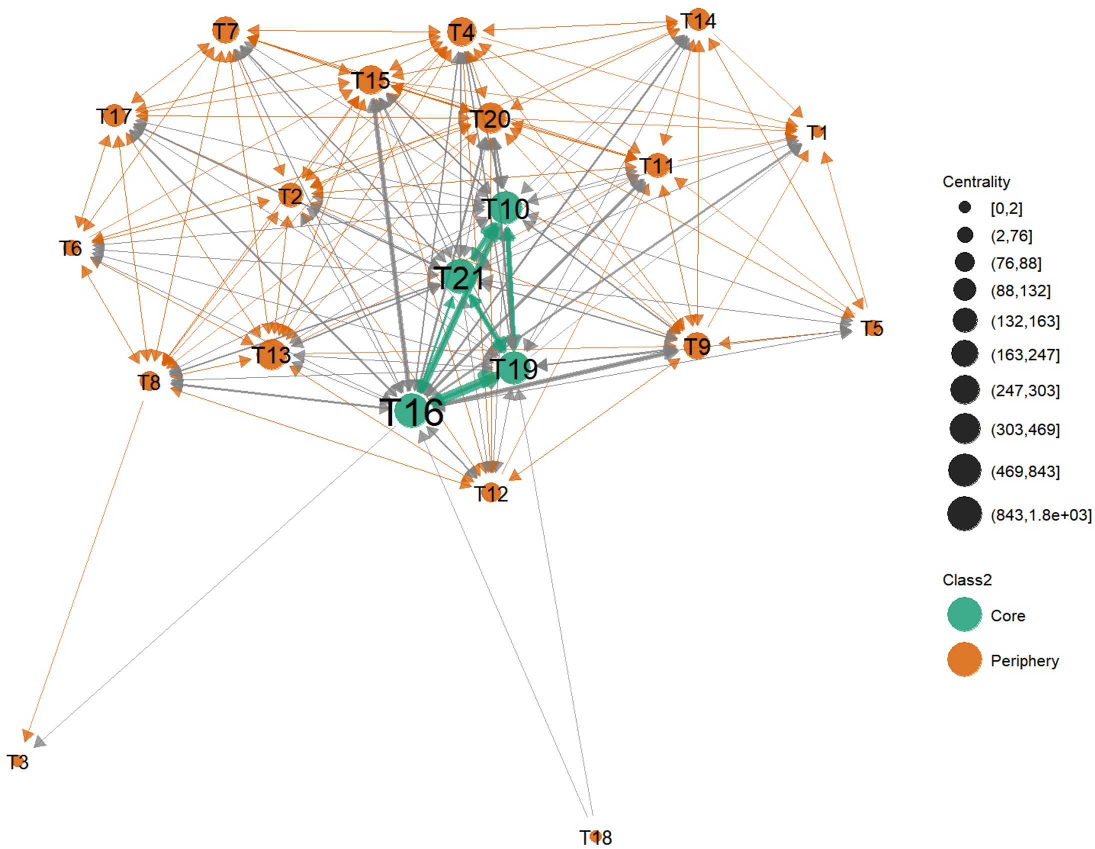

Figure 3: Forward BFOE core-periphery interaction

Note: This figure depicts the core-periphery structure of forward BFOE trading. Arrow directionality, edge weights, node and text sizes have the same meaning as in Figure 2. The color scheme represents the interaction of the *Core* and *Periphery*. An edge adopts the color of the node if the interaction is between nodes of the same group, or is gray for connections between nodes of different groups.

Core dealers are often 'making the market', as indicated by the relatively strong outgoing gray arrows to the periphery (see for example the edges from *T16* to *T8*, *T9*, *T11*, *T14*, *T15*, and *T17*), suggesting that the core traders are passively buying from or selling to the periphery. Given the functioning of Platts' eWindow, the core traders thereby reveal their intentions, as passive bids and offers have to be posted before the 16:25:00 cut-off for cash BFOE. Quotes can subsequently be amended until 16:30:00, and other traders can hit the bid or lift the offer of a market maker. Many thin gray edges target core traders (notice the concentration of gray arrows around the core nodes), suggesting that core traders also aggressively buy from or sell to a wide range of peripheral traders.

We hypothesize that the core-peripheral structure reflects the 'skin in the game' argument. The willingness and ability of traders to market make is closely linked to their business models and involvements in the upstream and downstream crude oil supply chains. More heavily invested traders have a better understanding of supply and demand levels (for example via their ownership or operation of oil fields and refineries) and have, therefore, better market making abilities and greater trading activities. This is then reflected in their centrality score. Hence, traders that are intricately involved in the physical trading of oil and often adopt the role of market makers are better informed about its fundamentals. The more central is a participant, the more telling is his trading activity for the financial oil market, leading to a price reaction from the futures market.

4.3 The impact of forward transactions on the futures market

This section tests the 'skin in the game' hypothesis and reports the main results of the study. Have transactions by more central forward traders a more pronounced impact on the price in the futures market? A likely source of price impact is fundamental supply and demand information, gained from involvement in upstream and downstream oil business lines, that is revealed to the futures market via forward trading.

To answer the research question, we compute the price impact of passively initiated forward buy and sell transactions on the futures market. This approach originates in the functioning of Platts' eWindow, where the so-called market makers reveal their intentions to buy or sell, as passive bids and offers have to be posted before the 16:25:00 cut-off for cash BFOE. Without this revelation of intentions, no trades will take place, as market takers can only aggressively hit or lift existing quotes. We are thus interested in the reaction of the futures market to the participants' divulged needs to buy or sell large quantities of crude oil. Transaction sizes in the forward market are very large (100,000 bbl) and comparable to equity block trades; We therefore adopt a similar methodology to the one established in that literature (see for example Anand et al., 2012; Chan and Lakonishok, 1993, 1995; Holthausen et al., 1987, 1990; Kraus and Stoll, 1972). We take every forward transaction and identify the futures price in the market at the time of the trade, as well as the futures prices before and after the trade.

The permanent effect (PE) is computed as

$$PE\,(\%) = ln\left(\frac{P_{post}}{P_{pre}}\right) * 100 \tag{12}$$

The total effect (TE) is defined as

$$TE\,(\%) = ln\left(\frac{P_{t}}{P_{pre}}\right) * 100 \tag{13}$$

Finally, we calculate the liquidity effect (LE) as

$$LE\,(\%) = ln\left(\frac{P_{t}}{P_{post}}\right) * 100 \tag{14}$$

where P_t is the futures price at the time, t, of the forward transaction. P_{pre} and P_{post} are the futures prices five minutes before and five minutes after the forward transaction respectively. We choose five-minute intervals because all forward transactions happen between 16:25:00 and 16:30:00, which is part of the Dated Brent benchmark assessment period, and we thus allow the futures price to adjust to the information introduced by physical OTC trading activity.[19]

In a second step we run the following regression specification:

$$DV_t = \alpha + \beta_1 CENT_i + \gamma' X_t + \epsilon_t \tag{15}$$

where DV_t is one of the three price impact measures (*PE*, *TE*, *LE*) assessing the effect of a forward transaction on the futures price. $CENT_i$ is the full-sample-period centrality of the forward trader i of the transaction in question, as explained in Section 4.2.[20] We follow the existing literature (see Aggarwal and Samwick, 1999; Li and Schürhoff, 2014; Milbourn, 2003) and use an empirical cumulative distribution function (ECDF) to normalize the weighted-outdegree centrality measure to the range [0 = least central; 1 = most central]. The ECDF transformation has the advantage of maintaining the original ordering of centrality and mitigating the biases introduced by skewness and outliers, while simplifying the economic interpretation of the centrality variable (Li and Schürhoff, 2014). As such, a one-unit increase in centrality corresponds to a trader improving from the least central, $CENT = 0$, to the most central, $CENT = 1$, position.[21] X_t is a vector of control variables explained in detail below and in Appendix A.2. Heteroskedasticity-robust standard errors are clustered by trader.[22]

Table 4 reports the results from estimating Equation 15 for buy and sell forward transactions and controlling for potential confounding effects. The results without controls can be found in Table 6 in Appendix A.3. The coefficient of interest is *CENT*, which indicates whether forward traders that are more central move the futures market more than other traders.

CENT in the first column shows that, with a one-unit increase in centrality, one would expect the permanent impact of a forward buy transaction on the futures price to rise significantly by 15 bps. Similarly, from the second column, a forward sell transaction by a participant with a one-unit higher centrality impacts the futures price significantly more, by an added -10 bps. The results suggest that the physical oil market contains information that is released via forward trading activity and subsequently incorporated into the futures price. Importantly, central market makers in the forward market seem to be more informed,

[19]Hence, P_{pre} and P_{post} fall outside of the 16:25:00 to 16:30:00 window. Moreover, our results are robust to choosing different window lengths such as 10 minutes and 15 minutes.

[20] We conduct robustness tests computing centrality on a yearly compounded basis. The unchanged results can be found in Appendix A.5.

[21]Applying a weighted ECDF, using the number of outgoing edges of a trader, does not materially affect the results.

[22]The results are unchanged if we cluster by date, maturity, and trader.

and therefore their trading activity has a larger price impact. Forward trader identities are visible to other market participants in the OTC trading setup of eWindow. The futures market appears to be alert to the identity of the trader and reacts more strongly to the actions of traders that are more central. This is in line with the literature on block trades (see for example Holthausen et al., 1987, 1990; Kraus and Stoll, 1972), and particularly the study by Chan and Lakonishok (1993), which recognizes trader identity as the dominant driver of price impact. The significant role played by forward market centrality in impacting the futures market price confirms the 'skin in the game' hypothesis.

We control for a variety of potentially confounding effects, without changing the insights obtained from our analysis. The control variables are the log futures volume over the price impact assessment window (*log(VOL)*), the standard deviation of futures log returns over the price impact assessment window (*log(VOLA)*), the forward buy volume in the front-month contract by trading day (*QBUY*), the forward sell volume in the front-month contract by trading day (*QSELL*), the log return between the forward transaction price at time t, and the first quote price of the related order ahead of execution (*log(PM)*), a dummy that takes the value 1 for companies that are oil majors and 0 otherwise (*OILM*), a dummy that takes the value 1 for companies that are commodity trading houses and 0 otherwise (*OILT*), the log Herfindahl-Hirschman Index by forward contract month, where the market share for each trader and contract month is determined by the gross notional of the forwards transacted (*log(HHI)*), a dummy that takes the value 1 after the 1 February 2015 to control for the potential effect of Platts changing the Dated Brent assessment period to 10-30 days ahead (*BMCHG*), a dummy that takes the value 1 after the 1 February 2016 to control for the potential effect of extending the expiry of the futures to two-months-ahead contract and thereby aligning it with the forward contract (*FUTCHG*), and, finally, the dummies accounting for day-of-the-week effects with Monday as the baseline category (*WEEKD()*).

For parsimony, we only discuss the implications for *PE*, the dependent variable of highest interest. On the one hand, *log(VOL)* does not affect the *PE* variable. On the other hand, in the event of a 1% change in *log(VOLA)*, the *PE* of buy and sell transactions is impacted significantly by -0.08% and -0.04% respectively. The *QBUY* on the day of the executed forward transaction has a statistically, although not economically, significant impact on both the buy and sell *PE*. The *QSELL* only significantly affects the permanent impact of a sell transaction. The price movement in the forward market ahead of the execution of a transaction (*log(PM)*) has a strong impact on the left-hand-side variable. A 1% change in the pre-execution forward price movement of a buy and sell transaction changes the *PE* by 14% and 12% respectively. The affiliation of the forward trader i to big oil (oil majors, *OILM*) or commodity trading (*OILT*) does not impact the coefficient of interest. The *log(HHI)* measuring market concentration and competition has a significant effect on the *PE* of both buy and sell forward transactions. A 1% change in the *log(HHI)* moves the buy and sell *PE* by -0.09% and -0.08% respectively. The dummy variables *BMCHG* and *FUTCHG*, controlling for changes in the forward and futures expiries respectively, do not affect the regression outcome.[23] Finally, day-of-the-week effects (*WEEKD(WED)* for buy and *WEEKD(FRI)* for sell trades) have a significant influence on *PE*. Overall, even after controlling for a variety of possibly interfering effects and events, the conclusions regarding centrality and its price impact remain unchanged.

The adjusted R^2 for the *PE* regressions is 8% for buys and 10% for sells. This is within the range of other studies analyzing the effects of network dynamics on trading variables; for instance, Di Maggio et al. (2017b) report R^2 values between 2% and 8%.

[23]It should be noted, however, that *BMCHG* and *log(VOLA)* have a Pearson correlation of 69% (see Appendix A.4), suggesting that futures volatility increased with the changes that were made to the forward contract. *log(VOLA)* might therefore already capture part of this effect. *BMCHG* and *FUTCHG* are also correlated by 56%.

Table 4: Price impact of forward trades on futures market: With controls

	Dependent variable:					
	PE		TE		LE	
	Buy	Sell	Buy	Sell	Buy	Sell
CENT	0.15^{***} (0.05)	-0.10^{***} (0.04)	0.20^{***} (0.06)	-0.14^{***} (0.03)	0.05 (0.04)	-0.04^{*} (0.02)
log(VOL)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	-0.01^{**} (0.01)
log(VOLA)	-0.08^{**} (0.03)	-0.04^{**} (0.02)	-0.08^{***} (0.02)	-0.06^{***} (0.02)	−0.00 (0.02)	−0.01 (0.02)
QBUY	0.00^{***} (0.00)	0.00^{***} (0.00)	0.00^{***} (0.00)	0.00 (0.00)	−0.00 (0.00)	-0.00^{***} (0.00)
QSELL	−0.00 (0.00)	-0.00^{***} (0.00)	-0.00^{**} (0.00)	-0.00^{***} (0.00)	−0.00 (0.00)	0.00 (0.00)
log(PM)	14.26^{***} (2.04)	11.75^{***} (1.39)	14.91^{***} (1.20)	12.31^{***} (1.34)	0.65 (1.40)	0.56 (0.53)
OILM	−0.00 (0.03)	0.03 (0.02)	−0.03 (0.03)	0.01 (0.02)	-0.03^{*} (0.02)	−0.02 (0.02)
OILT	0.00 (0.03)	0.02 (0.02)	−0.03 (0.03)	0.01 (0.02)	−0.04 (0.02)	−0.01 (0.01)
log(HHI)	-0.09^{***} (0.02)	-0.08^{***} (0.02)	-0.03^{***} (0.01)	−0.02 (0.03)	0.06^{***} (0.01)	0.05^{***} (0.02)
BMCHG	0.00 (0.05)	0.00 (0.02)	0.00 (0.03)	0.04^{*} (0.02)	0.00 (0.02)	0.03^{**} (0.02)
FUTCHG	0.05 (0.04)	−0.02 (0.03)	0.05^{*} (0.03)	−0.01 (0.03)	−0.01 (0.03)	0.02 (0.01)
WEEKD(TUE)	0.04 (0.03)	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)	−0.02 (0.02)	0.01 (0.01)
WEEKD(WED)	-0.06^{***} (0.02)	0.00 (0.02)	-0.04^{***} (0.01)	−0.01 (0.02)	0.02 (0.01)	−0.01 (0.02)
WEEKD(THU)	−0.02 (0.03)	−0.01 (0.02)	−0.03 (0.02)	−0.00 (0.02)	−0.00 (0.02)	0.00 (0.01)
WEEKD(FRI)	−0.02 (0.03)	-0.05^{**} (0.02)	−0.01 (0.02)	−0.00 (0.02)	0.02 (0.02)	0.04^{**} (0.02)
Constant	-1.15^{***} (0.36)	-0.44^{**} (0.20)	-1.03^{***} (0.23)	−0.33 (0.23)	0.13 (0.20)	0.11 (0.15)
Observations	2,083	2,473	2,083	2,473	2,083	2,473
R^2	0.09	0.10	0.15	0.12	0.03	0.04
Adjusted R^2	0.08	0.10	0.14	0.11	0.02	0.04
Residual Std. Error	0.27 (df = 2067)	0.26 (df = 2457)	0.21 (df = 2067)	0.22 (df = 2457)	0.18 (df = 2067)	0.17 (df = 2457)

Note: *p<0.1; **p<0.05; ***p<0.01

CENT measures the centrality of the forward market participants in terms of the ECDF-normalized weighted out-degree [0 = least central; 1 = most central]. Please refer to Appendix A.2 for a detailed explanation of the control variables. The coefficients are reported in percentage terms (%). Robust standard errors clustered at the trader level are reported in parentheses.

The results for the total price impact in the third and fourth columns align with those for the permanent price impact. A one-unit increase in forward trader centrality leads to a significantly stronger *TE* of forward buy transactions on the futures market, the increase being 20 bps. In the same vein, if a forward trader moves from least to most central, the sell transaction in the forward market impacts the futures market by a significant total of -14 bps. The adjusted R^2 for these regressions ranges from 11% to 14%.

Lastly, the liquidity effect, shown in the fifth column of Table 4, of a forward buy transaction on the futures price is insignificant. For the liquidity effect in the sixth column, we find that a one-unit rise in centrality leads to a significant reversal at the 10% level in the futures price—the *LE* of a forward sell transaction amounts to -4 bps. The adjusted R^2 here lies between 2% and 4%.

All in all, our findings support our 'skin in the game' hypothesis. Trading activity by central forward participants seems to convey valuable information to the financial market that is subsequently impounded into futures prices.

4.4 Robustness tests

In this section, we corroborate that it is indeed the centrality in the forward trading network that matters. As described in Section 2, other products are traded in the physical market during the Platts Window. The OTC-traded CFD market is the most liquid of those, while the cargo market is the least liquid, as measured by the number of trades and quotes. While the CFD and cash BFOE markets are closely interlinked, the participant groups of both markets are similar but different at the same time. For example, some participants who are very active in the forward market occupy a less prominent role in the CFD market and vice versa, and again others are crucial participants in both. Additionally, some engaged CFD traders decide not to participate in the forward market at all. At the same time, all forward traders participate in the CFD market. Hence, we compute the centrality of all traders in the CFD market and substitute the forward trader centrality used in the previous section with the CFD centrality, to determine the importance of the traders anew. CFD trading allows market participants to minimize the risk arising from price differentials between elements of the Brent complex, and therefore forward traders with high CFD centrality scores might be well informed about oil fundamentals too.

Table 5 shows that the CFD *CENT* coefficient is insignificant in explaining the *PE*, *TE*, and *LE* of forward transactions on the futures price.[24] This finding supports our assertion that the forward network centrality is a valuable proxy for 'skin in the game' information from upstream and downstream business lines. The fact that cash BFOE contracts are used to trade long-term supply and demand, while CFDs serve to manage short-term exposures and to hedge price risks of the Brent complex, might help to explain

[24] We also test the importance of the forward and CFD centrality measures in jointly explaining the price impact in the futures market. While the forward centrality is highly significant, the CFD centrality does not affect the price impact variables.

the difference in importance. In addition, forward trading requires the ability to receive and deliver physical oil, while CFDs are cash-settled derivatives (see Barret, 2012). The business of forward participants thus demands higher infrastructure investments and closer integration with the upstream and downstream petroleum industry. Given the closeness of forwards and futures, the link is stronger and the information is more easily observed and impounded. Therefore, forward network centrality is a valid proxy for supply and demand fundamentals in the physical oil market that are revealed via trading and subsequently incorporated into futures prices.

Table 5: Price impact of forward trades: CFD market centrality

	Dependent variable:					
	PE		TE		LE	
	Buy	Sell	Buy	Sell	Buy	Sell
CENT	0.07 (0.05)	−0.02 (0.03)	0.10 (0.07)	−0.03 (0.05)	0.03 (0.04)	−0.01 (0.04)
log(VOL)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	−0.01** (0.01)
log(VOLA)	−0.08** (0.03)	−0.05** (0.02)	−0.08*** (0.02)	−0.06*** (0.02)	−0.00 (0.02)	−0.02 (0.02)
QBUY	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00 (0.00)	−0.00 (0.00)	−0.00*** (0.00)
QSELL	−0.00 (0.00)	−0.00*** (0.00)	−0.00** (0.00)	−0.00*** (0.00)	−0.00 (0.00)	0.00 (0.00)
log(PM)	14.29*** (2.08)	11.67*** (1.38)	14.95*** (1.33)	12.22*** (1.29)	0.66 (1.40)	0.55 (0.54)
OILM	0.04 (0.03)	0.00 (0.02)	0.01 (0.03)	−0.03 (0.02)	−0.02 (0.02)	−0.03* (0.02)
OILT	0.03 (0.02)	−0.00 (0.02)	0.01 (0.03)	−0.02 (0.02)	−0.03 (0.02)	−0.01 (0.02)
log(HHI)	−0.09*** (0.02)	−0.08*** (0.02)	−0.02** (0.01)	−0.03 (0.03)	0.06*** (0.01)	0.05*** (0.02)
BMCHG	−0.00 (0.05)	0.01 (0.03)	−0.00 (0.03)	0.04* (0.02)	0.00 (0.02)	0.04** (0.02)
FUTCHG	0.05 (0.04)	−0.02 (0.03)	0.04 (0.03)	0.00 (0.03)	−0.01 (0.03)	0.02 (0.01)
WEEKD(TUE)	0.04 (0.03)	0.01 (0.02)	0.01 (0.02)	0.02 (0.02)	−0.02 (0.02)	0.01 (0.01)
WEEKD(WED)	−0.06*** (0.02)	0.01 (0.02)	−0.04*** (0.01)	−0.01 (0.02)	0.02 (0.01)	−0.01 (0.02)
WEEKD(THU)	−0.02 (0.03)	−0.01 (0.02)	−0.03 (0.02)	−0.00 (0.02)	−0.00 (0.02)	0.00 (0.01)
WEEKD(FRI)	−0.02 (0.03)	−0.04*** (0.02)	−0.00 (0.02)	−0.00 (0.02)	0.02 (0.02)	0.04** (0.02)
Constant	−1.10*** (0.37)	−0.51** (0.20)	−0.97*** (0.23)	−0.42** (0.21)	0.13 (0.18)	0.09 (0.16)
Observations	2,083	2,473	2,083	2,473	2,083	2,473
R^2	0.09	0.10	0.13	0.11	0.03	0.04
Adjusted R^2	0.08	0.09	0.13	0.10	0.02	0.04
Residual Std. Error	0.27 (df = 2067)	0.26 (df = 2457)	0.21 (df = 2067)	0.23 (df = 2457)	0.18 (df = 2067)	0.17 (df = 2457)

Note: *p<0.1; **p<0.05; ***p<0.01

CENT measures the physical CFD market trader centrality in terms of ECDF-normalized weighted out-degree [0 = least central; 1 = most central]. Please refer to Appendix A.2 for a detailed explanation of the control variables. The coefficients are reported in percentage terms (%). Robust standard errors clustered at the trader level are reported in parentheses.

5 Conclusion

Despite the fact that the financial and physical oil markets are, and have historically been, inextricably linked, our understanding of the futures market has gradually increased while we still know very little about its physical counterpart—the forward market.

We create a unique and novel dataset by combining intraday data for both markets. We confirm the longstanding belief that the futures market is nowadays the dominant information leader, incorporating the majority of new information ahead of the forward market. This finding is unsurprising given that the virtually 24-hour exchange-traded and financially settled futures contracts are by design more active. However, the forward market, with its proportionally fewer quotes and transactions and only a select number of active participants, contributes a non-trivial amount to oil price discovery. During only five minutes of active trading, from 16:25 to 16:30, at the end of the day, forwards impound approximately 20% of the innovations to the efficient price of oil. The forward price is also less noisy than the futures price. This is in line with the findings of Frino et al. (2017), suggesting that indeed physical market activity during the time of the Dated Brent benchmark assessment does indeed substantially influence the futures price development.

Lastly, we show that information from the physical market is revealed via forward trading and subsequently incorporated into futures prices. In support of our hypothesis, we find that more central forward participants with substantial 'skin in the game' have a more pronounced futures price impact. A one-unit increase in forward network centrality corresponds to a 10 bps to 15 bps stronger permanent price impact. The informational advantage of central traders likely stems from proprietary business insights gleaned from their oil supply chains, for example through infrastructure stakes, such as oil field or refinery ownership, and trading relationships with other major players in the market. The results suggest that fundamental supply and demand information is a significant driver of commodity prices.

Our findings need to be interpreted in the light of a few limitations. First, forward trading is limited to a very short period every day. We do not wish to make any inferences about oil price discovery outside of this window. Future research should aim to reconcile data on ET derivatives with that on other OTC derivatives and investigate their interactions. CFDs, for example, play a crucial role in the physical oil market too. Second, the data limitations that cause difficulties in the establishment of cointegration between oil futures and forwards on an intraday basis show there is a call for caution when interpreting the price discovery findings. While the results are conservative, the price discovery metrics depend, by design, on the specifications of the VECM.

Despite these constraints, we confirm assertions in the literature that the financialization of commodity markets substantially affects the way oil is traded (see Cheng and Xiong, 2014). However, we underline that there is a close interaction between financial and physical contracts, with unique features of both markets contributing to the determination of the efficient oil price.

References

Aggarwal, Rajesh K., and Andrew A. Samwick, 1999, The other side of the trade-off: The impact of risk on executive compensation, *Journal of Political Economy* 107, 65–105.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, 2012, Performance of institutional trading desks: An analysis of persistence in trading costs, *The Review of Financial Studies* 25, 557–598.

Baillie, Richard T., G. Geoffrey Booth, Yiuman Tse, and Tatyana Zabotina, 2002, Price discovery and common factor models, *Journal of Financial Markets* 5, 309–321.

Barret, Christophe, 2012, Brent prices: Impact of PRA methodology on price formation, Oxford Energy Comment, The Oxford Institute for Energy Studies.

Basak, Suleyman, and Anna Pavlova, 2016, A model of financialization of commodities, *The Journal of Finance* 71, 1511–1556.

Bekiros, Stelios D., and Cees G. H. Diks, 2008, The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality, *Energy Economics* 30, 2673–2685.

Booth, G. Geoffrey, Raymond W. So, and Yiuman Tse, 1999, Price discovery in the German equity index derivatives markets, *The Journal of Futures Markets* 19, 619–643.

Büyükşahin, Bahattin, and Jeffrey H. Harris, 2011, Do speculators drive crude oil futures prices?, *The Energy Journal* 32, 167–202.

Büyükşahin, Bahattin, and Michel A. Robe, 2014, Speculators, commodities and cross-market linkages, *Journal of International Money and Finance* 42, 38–70.

Chan, Louis K. C., and Josef Lakonishok, 1993, Institutional trades and intraday stock price behavior, *Journal of Financial Economics* 33, 173–199.

Chan, Louis K. C., and Josef Lakonishok, 1995, The behavior of stock prices around institutional trades, *The Journal of Finance* 50, 1147–1174.

Cheng, Ing-Haw, and Wei Xiong, 2014, Financialization of commodity markets, *Annual Review of Financial Economics* 6, 419–441.

Cheong, Serene, Dan Murtaugh, and Sharon Cho, 2017, Commodity traders have a really big problem, Bloomberg, `https://www.bloomberg.com/news/articles/2017-05-21/commodity-traders-lament-world-where-everybody-knows-everything`, [Online; accessed 2018-02-07].

Chu, Quentin C., Wen-Liang G. Hsieh, and Yiuman Tse, 1999, Price discovery on the S&P 500 index markets: An analysis of spot index, index futures, and SPDRs, *International Review of Financial Analysis* 8, 21–34.

Cifarelli, Giulio, and Giovanna Paladino, 2010, Oil price dynamics and speculation: A multivariate financial approach, *Energy Economics* 32, 363–372.

Davis, Mike, 2012, Oil price benchmarks in international trade - Benchmark pricing: A co-dependent matrix, *Oxford Energy Forum* 87, 14–17.

Di Maggio, Marco, Francesco Franzoni, Amir Kermani, and Carlo Sommavilla, 2017a, The relevance of broker networks for information diffusion in the stock market, NBER Working Paper No. w23522, National Bureau of Economic Research.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, 2017b, The value of trading relations in turbulent times, *Journal of Financial Economics* 124, 266–284.

Easley, David, and Jon Kleinberg, 2010, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World* (Cambridge University Press, New York).

Fattouh, Bassam, 2011, An anatomy of the crude oil pricing system, Oxford Energy Comment WPM 40, The Oxford Institute for Energy Studies.

Fattouh, Bassam, Lutz Kilian, and Lavan Mahadeva, 2013, The role of speculation in oil markets: What have we learned so far?, *The Energy Journal* 34, 7–33.

Figuerola-Ferretti, Isabel, and Jesús Gonzalo, 2010, Modelling and measuring price discovery in commodity markets, *Journal of Econometrics* 158, 95–107.

Frino, Alex, Gbenga Ibikunle, Vito Mollica, and Tom Steffen, 2017, The impact of commodity benchmarks on derivatives markets: The case of the Dated Brent assessment and Brent futures, *Journal of Banking & Finance* forthcoming.

Frino, Alex, Andrew Lepone, Vito Mollica, and Shunquan Zhang, 2016, Are hedgers informed? An examination of the price impact of large trades in illiquid agricultural futures markets, *The Journal of Futures Markets* 36, 612–622.

Garbade, Kenneth D., and William L. Silber, 1983, Price movements and price discovery in futures and cash markets, *The Review of Economics and Statistics* 65, 289–297.

Gonzalo, Jesus, and Clive Granger, 1995, Estimation of common long-memory components in cointegrated systems, *Journal of Business & Economic Statistics* 13, 27–35.

Goyal, Sanjeev, 2005, Learning in networks, in Gabrielle Demange, and Myrna Wooders, eds., *Group Formation in Economics: Networks, Clubs, and Coalitions*, 122–168 (Cambridge University Press, Cambridge).

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American Economic Review* 70, 393–408.

Hamilton, James D., 2009, Understanding crude oil prices, *The Energy Journal* 30, 179–206.

Hamilton, James D., and Jing C. Wu, 2015, Effects of index-fund investing on commodity futures prices, *International Economic Review* 56, 187–205.

Harris, Frederick H. de B., Thomas H. McInish, and Robert A. Wood, 2002, Security price adjustment across exchanges: An investigation of common factor components for Dow stocks, *Journal of Financial Markets* 5, 277–308.

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

Hasbrouck, Joel, 2003, Intraday price formation in U.S. equity index markets, *The Journal of Finance* 58, 2375–2400.

Hasbrouck, Joel, and George Sofianos, 1993, The trades of market makers: An empirical analysis of NYSE specialists, *The Journal of Finance* 48, 1565–1593.

Hellwig, Martin F., 1980, On the aggregation of information in competitive markets, *Journal of Economic Theory* 22, 477–498.

Hendershott, Terrence, Dan Li, Dmitry Livdan, and Norman Schürhoff, 2017, Relationship trading in OTC markets, SFI Research Paper No. 17-30, Swiss Finance Institute.

Henderson, Brian J., Neil D. Pearson, and Li Wang, 2015, New evidence on the financialization of commodity markets, *The Review of Financial Studies* 28, 1285–1311.

Holthausen, Robert W., Richard W. Leftwich, and David Mayers, 1987, The effect of large block transactions on security prices: A cross-sectional analysis, *Journal of Financial Economics* 19, 237–267.

Holthausen, Robert W., Richard W. Leftwich, and David Mayers, 1990, Large-block transactions, the speed of response, and temporary and permanent stock-price effects, *Journal of Financial Economics* 26, 71–95.

Inci, A. Can, and H. Nejat Seyhun, 2017, Degree of integration between Brent oil spot and futures markets: Intraday evidence, *Emerging Markets Finance and Trade* forthcoming.

Irwin, Scott H., and Dwight R. Sanders, 2011, Index funds, financialization, and commodity futures markets, *Applied Economic Perspectives and Policy* 33, 1–31.

Irwin, Scott H., Dwight R. Sanders, and Robert P. Merrin, 2009, Devil or angel? The role of speculation in the recent commodity price boom (and bust), *Journal of Agricultural and Applied Economics* 41, 377–391.

Jackson, Matthew O., 2005, A survey of network formation models: Stability and efficiency, in Gabrielle Demange, and Myrna Wooders, eds., *Group Formation in Economics: Networks, Clubs, and Coalitions*, 11–57 (Cambridge University Press, Cambridge).

Jackson, Matthew O., 2008, *Social and Economic Networks* (Princeton University Press, Princeton, N.J.).

Johansen, Søren, 1995, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Advanced texts in econometrics (Oxford University Press, Oxford).

Juvenal, Luciana, and Ivan Petrella, 2015, Speculation in the oil market, *Journal of Applied Econometrics* 30, 621–649.

Kaufmann, Robert K., and Ben Ullman, 2009, Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices, *Energy Economics* 31, 550–558.

Kilian, Lutz, 2009, Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market, *The American Economic Review* 99, 1053–1069.

Kilian, Lutz, and Daniel P. Murphy, 2014, The role of inventories and speculative trading in the global market for crude oil, *Journal of Applied Econometrics* 29, 454–478.

Knittel, Christopher R., and Robert S. Pindyck, 2016, The simple economics of commodity price speculation, *American Economic Journal: Macroeconomics* 8, 85–110.

Kraus, Alan, and Hans R. Stoll, 1972, Price impacts of block trading on the New York Stock Exchange, *The Journal of Finance* 27, 569–588.

Lehmann, Bruce N., 2002, Some desiderata for the measurement of price discovery across markets, *Journal of Financial Markets* 5, 259–276.

Li, Dan, and Norman Schürhoff, 2014, Dealer Networks, SFI Research Paper No. 14-50, Swiss Finance Institute.

Liu, Wai-Man, Emma Schultz, and John Swieringa, 2015, Price dynamics in global crude oil markets, *The Journal of Futures Markets* 35, 148–162.

Milbourn, Todd T., 2003, CEO reputation and stock-based compensation, *Journal of Financial Economics* 68, 233–262.

Munyan, Benjamin, and Sumudu Watugala, 2017, Market dominance in bond and CDS interdealer networks, Working Paper, Vanderbilt University and Cornell University.

Ozsoylev, Han N., Johan Walden, M. Deniz Yavuz, and Recep Bildik, 2014, Investor networks in the stock market, *The Review of Financial Studies* 27, 1323–1366.

Putniņš, Tālis J., 2013, What do price discovery metrics really measure?, *Journal of Empirical Finance* 23, 68–83.

Quan, Jing, 1992, Two-step testing procedure for price discovery role of futures prices, *The Journal of Futures Markets* 12, 139–149.

Schwarz, Thomas V., and Andrew C. Szakmary, 1994, Price discovery in petroleum markets: Arbitrage, cointegration, and the time interval of analysis, *The Journal of Futures Markets* 14, 147–167.

Silvapulle, Param, and Imad A. Moosa, 1999, The relationship between spot and futures prices: Evidence from the crude oil market, *The Journal of Futures Markets* 19, 175–193.

Silvennoinen, Annastiina, and Susan Thorp, 2013, Financialization, crisis and commodity correlation dynamics, *Journal of International Financial Markets, Institutions and Money* 24, 42–65.

Singleton, Kenneth J., 2013, Investor flows and the 2008 boom/bust in oil prices, *Management Science* 60, 300–318.

Sockin, Michael, and Wei Xiong, 2015, Informational frictions and commodity markets, *The Journal of Finance* 70, 2063–2098.

Stock, James H., and Mark W. Watson, 1988, Testing for common trends, *Journal of the American Statistical Association* 83, 1097–1107.

Tang, Ke, and Wei Xiong, 2012, Index investment and the financialization of commodities, *Financial Analysts Journal* 68, 54–74.

Tse, Yiuman, 2000, Further examination of price discovery on the NYSE and regional exchanges, *The Journal of Financial Research* 23, 331–351.

Yan, Bingcheng, and Eric Zivot, 2010, A structural analysis of price discovery measures, *Journal of Financial Markets* 13, 1–19.

A Appendix

A.1 Data filters

For the price discovery analysis:

- We only include the forward front-month and the respective futures front-month activity.
- We exclude observations where bid > ask and where either the bid, ask, or last trade price equals zero and omit forward quotes that are more than four standard deviations away from the daily mean.
- We exclude days where we are unable to compute a forward mid-price because either bid or ask quotes are unavailable over the whole trading day.
- We only include forward trading from 16:22:00 to 16:30:00 (on normal Platts Dated Brent assessment days) and 12:22:00 to 12:30:00 (on early Platts Dated Brent assessment days) respectively. The reason is that, due to the functioning of the Platts eWindow, 99% of the activity takes place in this time window at the end of each trading day. The early assessment days are 2012-04-05, 2012-12-24, 2012-12-31, 2013-03-28, 2013-12-24, 2013-12-31, 2014-04-17, 2014-12-24, 2014-12-31, 2015-04-02, 2015-12-24, 2015-12-31, 2016-03-24, 2016-12-23, 2016-12-30, 2017-04-13, 2017-12-22, and 2017-12-29.
- On each trading day, for the calculation of the price discovery metrics, we define the first timestamp to be the time $t_1 = 1$ of the first forward quote ($n = 1$). The last timestamp corresponds to the time 60 seconds after the last forward quote on that day ($n = N$), $t_N = T + 60$, where $n = 1, ..., N$ and $t = 1, ..., T$. We thus do a full join of futures and forward data within the time range $[t_1 = 1; t_N = T + 60]$. We do this to allow for a potential adjustment of the futures price to the last forward quote. Doing this, we also avoid using a standardized time window and thereby biasing our results by including many stale forward quotes. For example, on one day the last forward quote might be received at 16:27:35, while on the next day the last forward quote might only arrive at 16:29:58. If we were to sample on a fixed window, we would, in the first case, include a vast number of futures quotes up to 16:30:00, after price discovery in the forward market had already stopped, and the forward quote would be stale for more than two minutes.
- We remove all days on which the variation in the forward quotes is below the first percentile level of the quote variation on all days. We need a minimum quote variation in the forward market to establish cointegration.
- We remove days on which the forward and futures contracts are not cointegrated at the 75% confidence level or higher. We use the AIC test to determine the optimal number of lags, allowing for a maximum of 60 lags.

For the price impact analysis:

- We only include the forward front-month and the respective futures front-month activity.
- We exclude observations where bid > ask and where either the bid, ask, or last trade price equals zero and exclude forward quotes that are more than four standard deviations away from the daily mean.
- We only include front-month forward transactions, all of which were executed between 16:25:00 and 16:30:00 (on normal Platts Dated Brent assessment days) and 12:25:00 and 12:30:00 (on early Platts Dated Brent assessment days) respectively.
- We compute the price impact measures on the futures price over a 10-minute window, using $+-$five minutes to determine the pre- and post-benchmark prices. Our results are robust to choosing $+-10$ minutes or $+-15$ minutes instead. We do not use less than five minutes because, for a robust calculation of the price impact, the pre- and post-benchmark prices should fall outside the 16:25:00–16:30:00 period of the Platts Dated Brent assessment.

A.2 Control variables

- *CENT* is the ECDF-normalized weighted out-degree measure of each forward trader, ranging from 0 = least central to 1 = most central, calculated from the network presented in Section 4.2. The network and centrality are determined based on all forward transactions in all contract months over the full sample period. The reason for using the full sample period is that we aim to measure the importance of a market maker, and its reputation as a major trading participant, established over time. We use transactions in all contract months to capture the overall standing of a trader in the market, even though we only measure the front-month price impact.
- *log(VOL)* is the log futures volume over the price impact assessment window.
- *log(VOLA)* is the standard deviation of futures log returns over the price impact assessment window.
- *QBUY* is the forward buy volume in the front-month contract by trading day.
- *QSELL* is the forward sell volume in the front-month contract by trading day.
- *log(PM)* is the log return between the forward transaction price at time t, and the first quote price of the related order ahead of execution. This variable accounts for potential price adjustments in the forward market in the direction of the upcoming trade, ahead of its completion.
- *OILM* is a dummy that takes the value 1 for companies that are oil majors and 0 otherwise.

- *OILT* is a dummy that takes the value 1 for companies that are commodity trading houses and 0 otherwise.
- *log(HHI)* is the log Herfindahl-Hirschman Index by forward contract month, where the market share for each trader and contract month is determined by the gross notional of the forwards transacted. This variable approximates and controls for market concentration and competition.
- *BMCHG* takes the value 1 after 2015-02-01 to control for the potential effect of Platts changing the Dated Brent assessment period to 10-30 days ahead. This had an impact on the expiry of BFOE forwards too.
- *FUTCHG* takes the value 1 after 2016-02-01 to control for the potential effect of changes to the futures contract expiry, extending it to a two-months-ahead contract and thereby aligning it with the forward contract.
- *WEEKD()* are dummies accounting for day-of-the-week effects. The baseline category is Monday.

A.3 Price impact regressions: Without controls

Table 6: Price impact of forward trades on futures market: Without controls

	Dependent variable:					
	PE		TE		LE	
	Buy	Sell	Buy	Sell	Buy	Sell
CENT	0.12*** (0.03)	−0.12*** (0.04)	0.16*** (0.05)	−0.16*** (0.03)	0.04 (0.04)	−0.03 (0.02)
Constant	−0.08*** (0.03)	0.05** (0.03)	−0.12*** (0.04)	0.07*** (0.02)	−0.04 (0.03)	0.02 (0.02)
Observations	2,083	2,473	2,083	2,473	2,083	2,473
R^2	0.01	0.01	0.02	0.02	0.00	0.00
Adjusted R^2	0.01	0.01	0.02	0.02	0.00	0.00
Residual Std. Error	0.28 (df = 2081)	0.27 (df = 2471)	0.22 (df = 2081)	0.24 (df = 2471)	0.18 (df = 2081)	0.18 (df = 2471)

Note: *p<0.1; **p<0.05; ***p<0.01

CENT measures the centrality of the forward market participants in terms of the ECDF-normalized weighted out-degree [0 = least central; 1 = most central]. Please refer to Appendix A.2 for a detailed explanation of the control variables. The coefficients are reported in percentage terms (%). Robust standard errors clustered at the trader level are reported in parentheses.

A.4 Correlation matrix

Table 7: Correlation matrix of control variables

	CENT	log(VOL)	log(VOLA)	QBUY	QSELL	log(PM)	OILM	OILT	log(HHI)	BMCHG	FUTCHG	MON	TUE	WED	THU	FRI
CENT	1	-0.03	-0.07	-0.00	-0.04	0.05	0.43	-0.08	0.16	-0.13	-0.18	-0.00	-0.02	-0.02	0.02	0.02
log(VOL)	-0.03	1	0.31	-0.05	0.04	-0.01	-0.03	0.03	-0.07	0.17	0.20	-0.04	0.05	0.04	-0.09	0.03
log(VOLA)	-0.07	0.31	1	-0.09	0.21	-0.11	-0.12	0.12	-0.34	0.69	0.38	-0.04	0.02	0.08	-0.03	-0.04
QBUY	-0.00	-0.05	-0.09	1	-0.27	0.26	0.03	-0.01	0.04	-0.11	-0.02	0.02	-0.06	0.06	0.03	-0.05
QSELL	-0.04	0.04	0.21	-0.27	1	-0.26	-0.11	0.08	-0.14	0.16	-0.02	0.01	0.07	0.03	-0.02	-0.09
log(PM)	0.05	-0.01	-0.11	0.26	-0.26	1	0.06	-0.05	0.06	-0.09	0.04	-0.00	-0.00	-0.00	0.00	0.01
OILM	0.43	-0.03	-0.12	0.03	-0.11	0.06	1	-0.76	0.16	-0.12	-0.12	-0.02	0.01	-0.02	0.00	0.04
OILT	-0.08	0.03	0.12	-0.01	0.08	-0.05	-0.76	1	-0.13	0.14	0.09	0.03	-0.04	0.02	0.03	-0.04
log(HHI)	0.16	-0.07	-0.34	0.04	-0.14	0.06	0.16	-0.13	1	-0.41	-0.16	0.00	-0.01	0.02	-0.02	0.01
BMCHG	-0.13	0.17	0.69	-0.11	0.16	-0.09	-0.12	0.14	-0.41	1	0.56	-0.02	0.04	-0.01	0.02	-0.04
FUTCHG	-0.18	0.20	0.38	-0.02	-0.02	0.04	-0.12	0.09	-0.16	0.56	1	-0.00	0.02	-0.00	0.02	-0.03
MON	-0.00	-0.04	-0.04	0.02	0.01	-0.00	-0.02	0.03	0.00	-0.02	-0.00	1	-0.27	-0.26	-0.25	-0.23
TUE	-0.02	0.05	0.02	-0.06	0.07	-0.00	0.01	-0.04	-0.01	0.04	0.02	-0.27	1	-0.27	-0.26	-0.24
WED	-0.02	0.04	0.08	0.06	0.03	-0.00	-0.02	0.02	0.02	-0.01	-0.00	-0.26	-0.27	1	-0.25	-0.24
THU	0.02	-0.09	-0.03	0.03	-0.02	0.00	0.00	0.03	-0.02	0.02	0.02	-0.25	-0.26	-0.25	1	-0.22
FRI	0.02	0.03	-0.04	-0.05	-0.09	0.01	0.04	-0.04	0.01	-0.04	-0.03	-0.23	-0.24	-0.24	-0.22	1

This table reports the Pearson correlation of the regression control variables. Please refer to Appendix A.2 for a detailed explanation of the control variables.

A.5 Additional results: Yearly compounded centrality

Table 8: Price impact of forward trades: Yearly compounded centrality

	Dependent variable:					
	PE		TE		LE	
	Buy	Sell	Buy	Sell	Buy	Sell
CENT	0.10^{***} (0.03)	-0.08^{**} (0.03)	0.12^{***} (0.05)	-0.12^{***} (0.03)	0.02 (0.03)	-0.04^{**} (0.02)
log(VOL)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	-0.01^{**} (0.01)
log(VOLA)	-0.08^{**} (0.03)	-0.04^{**} (0.02)	-0.08^{***} (0.02)	-0.06^{***} (0.02)	−0.00 (0.02)	−0.02 (0.02)
QBUY	0.00^{***} (0.00)	0.00^{***} (0.00)	0.00^{***} (0.00)	0.00 (0.00)	−0.00 (0.00)	-0.00^{***} (0.00)
QSELL	−0.00 (0.00)	-0.00^{***} (0.00)	-0.00^{**} (0.00)	-0.00^{***} (0.00)	−0.00 (0.00)	0.00 (0.00)
log(PM)	14.38^{***} (2.02)	11.85^{***} (1.39)	15.04^{***} (1.23)	12.46^{***} (1.33)	0.66 (1.41)	0.60 (0.55)
OILM	0.02 (0.02)	0.02 (0.02)	0.00 (0.02)	−0.00 (0.02)	−0.02 (0.02)	−0.02 (0.02)
OILT	0.02 (0.02)	0.01 (0.02)	−0.00 (0.02)	0.00 (0.02)	−0.03 (0.02)	−0.01 (0.01)
log(HHI)	-0.10^{***} (0.02)	-0.08^{***} (0.02)	-0.03^{***} (0.01)	−0.02 (0.03)	0.06^{***} (0.01)	0.05^{***} (0.02)
BMCHG	−0.00 (0.05)	0.00 (0.02)	0.00 (0.03)	0.04^{*} (0.02)	0.00 (0.02)	0.03^{**} (0.01)
FUTCHG	0.05 (0.04)	−0.02 (0.03)	0.04^{*} (0.03)	−0.01 (0.03)	−0.01 (0.03)	0.02 (0.01)
WEEKD(TUE)	0.03 (0.03)	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)	−0.02 (0.02)	0.01 (0.01)
WEEKD(WED)	-0.06^{***} (0.02)	0.00 (0.02)	-0.04^{***} (0.01)	−0.01 (0.02)	0.02 (0.01)	−0.01 (0.02)
WEEKD(THU)	−0.02 (0.03)	−0.00 (0.02)	−0.03 (0.02)	−0.00 (0.02)	−0.00 (0.02)	0.00 (0.01)
WEEKD(FRI)	−0.02 (0.03)	-0.04^{**} (0.02)	−0.00 (0.02)	−0.00 (0.02)	0.02 (0.02)	0.04^{**} (0.02)
Constant	-1.13^{***} (0.36)	-0.45^{**} (0.19)	-0.99^{***} (0.22)	−0.34 (0.23)	0.15 (0.19)	0.11 (0.16)
Observations	2,083	2,473	2,083	2,473	2,083	2,473
R^2	0.09	0.10	0.14	0.12	0.03	0.04
Adjusted R^2	0.08	0.10	0.13	0.11	0.02	0.04
Residual Std. Error	0.27 (df = 2067)	0.26 (df = 2457)	0.21 (df = 2067)	0.22 (df = 2457)	0.18 (df = 2067)	0.17 (df = 2457)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

CENT measures the yearly compounded trader centrality in terms of ECDF-normalized weighted out-degree [0 = least central; 1 = most central] and is computed starting with all trades from 2012-2013, then from 2012-2014, etc., until we incorporate all trades from 2012-2017. This allows us to account for changes in ranking over time and additions and withdrawals of participants. The coefficients are reported in percentage terms (%). Robust standard errors clustered at the trader level are reported in parentheses.