City Research Online

City, University of London Institutional Repository

Citation: de Menezes, L. M., Russo, M. & Urga, G. (2016). Identifying Drivers of Liquidity in the NBP Month-ahead Market. Paper presented at the EcoMod2016, 06 Jul 2016 - 08 Jul 2016, Lisbon, Portugal.

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: http://openaccess.city.ac.uk/17017/

Link to published version:

Copyright and reuse: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Identifying Drivers of Liquidity in the NBP Month-ahead Market

Marianna Russo*, Lilian M. de Menezes†, Giovanni Urga‡

29 April, 2016

Abstract

This study investigates the associations between measures of market quality such as liquidity, trading activity and volatility in the one-month-ahead forward market of the UK's National Balancing Point (NBP), from May 2010 to December 2014. The period of analysis includes when the EU Regulation on Market Integrity and Transparency (REMIT) came into force, hence, whether there were changes in the associations between those measures that could reflect REMIT is also investigated. Consistent with microstructure theory, a positive association between trading activity and volatility plus a negative association between volatility and subsequent liquidity are found. This is in line with the argument that in times of high trading activity and volatility, spread rises and market depth deteriorates, with implications for the costs of hedging and investment decisions. The results imply time-varying associations, but no significant differences were found in the dynamic of correlations between trading activity, volatility and liquidity after REMIT. From an energy policy perspective, the extent to which changes in liquidity might further nourish the effects of price fluctuations has implications for the competitiveness and efficiency of European natural gas markets.

Keywords: European natural gas markets, liquidity, market microstructure

1 Introduction

The liberalization of the natural gas market in the European Union has fostered the development of hubs and gas-to-gas competition. With increasing spot trading, there has been a progressive shift in pricing mechanisms from the traditional oil-linked pricing towards hub-linked pricing. In 2014, the share of gas traded that was indexed to hubs reached 61%, which shows a significant increase when compared to 15% in 2005 and 36% in 2010 (International Gas Union, 2015).

As energy companies respond to their increasing exposure to the spot market, a greater use of forward contracts is observed to hedge the increased price risk exposure (Pilipovic, 2007). Price volatility encourages the participation of investors, financial institutions and other non-physical traders, who further contribute to the development of trading hubs. Yet, concerns over the impact of investors on market quality have been raised by academic and other sources, mainly with respect to the trading in the less transparent over-the-counter (OTC) markets (Madhavan et al., 2005; Larosière, 2009; European Commission, 2011; Nijman, 2012), where trading can be customized and may not require collaterals.

In this context, liquidity, which is defined as the ability to match buyers and sellers at the lowest transaction cost (O'Hara, 1995) is of interest for researchers, market participants and policy makers. Liquidity

*Faculty of Finance, Cass Business School - City University London. 106, Bunhill Row, London

†Faculty of Management, Cass Business School - City University London. 106, Bunhill Row, London

‡Faculty of Finance and Centre for Econometric Analysis, Cass Business School - City University London. 106, Bunhill Row, London; University of Bergamo, Italy

affects the cost of hedging and investment decisions and is a barometer of market quality. The importance of liquidity is highlighted in the literature on market microstructure, which however has been mainly focused on financial assets, such as stocks and bonds (e.g. Chordia et al., 2000, 2005), or foreign exchanges (e.g. Bessembinder, 1994; Banti et al., 2012; Danielsson and Payne, 2012).

Overall, market microstructure literature has shown that low liquidity leads to lower asset prices and higher rate of returns, which are required to compensate investors for bearing the cost of liquidity (Amihud and Mendelson, 1986; Brennan and Subrahmanyam, 1996; Pastor and Stambaugh, 2003; Hasbrouck, 2009). Consequently, liquidity can be regarded as the ability of a market to offer sufficient opportunities for trading, so that any individual trade will have limited impact on market prices. A lack of liquidity may impede trading, thereby facilitating market concentration, and consequently has implications for market efficiency.

To date, studies of liquidity in energy markets remain rare. The present study therefore examines potential drivers of liquidity in the National Balancing Point (NBP) forward market, which is the main pricing hub in Europe (Cummins and Murphy, 2015; European Commission, 2015) and thus the best representative of the European natural gas market. In particular, the associations between liquidity, trading activity and price volatility are assessed through a vector autoregressive (VAR) representation, as in previous literature on financial markets (Chordia et al., 2005; Danielsson and Payne, 2012). In contrast to these studies, a time-varying approach is adopted, so that any change in the associations can be identified and further investigated. The period from May 2010 to December 2014 is considered, and includes when the Regulation (EU) No. 1227/2011 on wholesale Energy Market Integrity and Transparency (REMIT) came into force. REMIT has been in force since December 2011 but effective from 7 October 2015. It is however reasonable to expect that market players would have gradually started their preparation for the new rules, and that these may have impacted the associations that are examined.

REMIT implies that regulators must have frequent and timely access to records of transactions, as well as data on capacity, production, storage, consumption and transmission of electricity or natural gas. All participants in the liberalised European wholesale energy markets are required to provide very comprehensive data to the Agency for the Cooperation of Energy Regulators, in order to facilitate national and European regulators to ensure that prices are fair. Although higher transparency can improve liquidity, by reducing transaction costs and lowering barriers to market entry (European Commission, 2004; Bessembinder et al., 2013), REMIT's effects on liquidity and other aspects of market quality are unknown (Nijman, 2012).

The remainder of this study is organized as follows. In section 2, the literature on market microstructure theory is reviewed, and the new regulatory framework implied by REMIT is summarized. In section 3, data and methods are described. The empirical results are reported in section 4. Section 5 discusses the main findings. Finally, in section 6 implications and conclusions are drawn.

2 Liquidity in evolving market structures

2.1 Liquidity and the implications of the market microstructure theory

The literature on market microstructure analyzes how trading affects asset pricing, by addressing the consequences of buying and selling assets under explicit mechanisms and focusing on what O'Hara (1995) called the "dark side" of liquidity (p. 216). For example, costs of liquidity may be imposed on market players, and some investors may exit the market, thus creating instability and barriers to potential new entrants. Hence, it is important to understand the dynamics of liquidity in a market (Chordia et al., 2005).

According to Kyle (1985), liquidity summarizes the transactional properties of a market: tightness, i.e. "the cost of turning around a position over a short period of time"; depth, "the size of an order flow innovation required to change price of a given amount"; and resiliency, "the speed with which prices recover from a random, uninformative shock" (p. 1316). Together these properties highlight how the exchange between buyers and sellers actually occurs at any point in time.

The founder of market microstructure theory, Garman (1976), argued that exchange entails a flow of orders to buy and sell that may generate temporal imbalances between demand and supply, which affect the dynamics of liquidity over time. Consequently, the role of inventory has been analyzed. Amihud and Mendelson (1980) concluded that bid and ask prices depend on changes in the dealer's inventory positions, and thus are decreasing functions of her inventory imbalances. This argument led to the problem of the risk faced by the dealer in optimizing her inventory level. In Stoll (1978)'s view, the dealer provides a service in the form of immediacy supply and must be compensated. Accordingly, the cost of immediacy is the sum of: (1) holding costs, i.e. the price risk and opportunity cost of holding securities; (2) order costs, the costs of arranging, recording and clearing transactions; and (3) information costs, which arise if traders have superior information which adversely affect the dealer's expected returns. These transaction costs determine the bid-ask spread; in particular, holding costs guarantee the dealer's expected utility in spite of transactions that tend to move her away from the optimal inventory level, and may depend on the order flow, defined as the difference between buy and market sell volume, and the traded asset return volatility.

The informational-based approach to market microstructure, (Bagehot, 1971), relies on the theory of adverse selection to explain the bid-ask spread, which "reflects a balancing of losses to the informed with gains from the uninformed," (O'Hara, 1995, p. 54). In this context, the dealer's problem reduces to the optimization of gains or losses in a dynamic perspective where order flow is not exogenous but conveys information. That is, trading activity from informed traders represents the way in which information is spread in a market, or how uninformed traders can infer the underlying information on the asset value. Accordingly, order flow and trading activity are "signals" of information (Glosten and Milgrom, 1985; Easley and O'Hara, 1987), and asset pricing is no longer independent of private information on the fair asset value, impounded in the order flow. Therefore, order flow affects asset prices over time and prices are not independent of past trading activity. The bid-ask spread, thus, reflects this dynamic trading mechanic of price discovery and is a compensation for trading with better-informed traders.

Overall, inventory- and informational-based approaches to market microstructure imply that high trading activity can reduce market liquidity temporarily (inventory cost) and may move asset prices permanently (informational cost). In conclusion, according to the market microstructure theory, co-movements in trading activity, asset returns volatility and liquidity should be analyzed. For example, in the early literature on liquidity in financial markets, (Benston and Hagerman, 1974; Stoll, 1978) volatility and order flow were assumed to determine liquidity. The idea behind this research is that the higher the asset return volatility, the higher the inventory risk, the lower is liquidity. This is reflected in the bid-ask prices, which in turn depend on the order flow. Notwithstanding, liquidity was shown to influence equilibrium stock prices and expected returns (e.g. Brennan and Subrahmanyam, 1996; Brennan et al., 1998; Amihud, 2002), while both liquidity and asset returns was found to be affected by order flow and imbalances in the stock markets (e.g. Chordia et al., 2002). Using a vector autoregressive (VAR) representation, Chordia et al. (2005) found stock and bond market liquidity to be driven by returns and their volatility, as well as order imbalances. A positive association between liquidity and returns in stock markets was also reported by Hasbrouck (1991) in a VAR setting, and Hasbrouck (2009) using a Bayesian Gibbs approach. In the FX market, similar evidence was provided by Bessembinder (1994), who used different measures of spread and price impacts, and by Danielsson and Payne (2012) through a VAR representation. Order flow was used, for instance, by Evans and Lyons (2002) as a measure of liquidity and was associated with the dynamics of asset pricing. Nevertheless, when academic studies of energy markets are considered, the assessment of associations across liquidity, return volatility and trading activity appears to have been neglected.

2.2 Regulation on Market Integrity and Transparency (REMIT) and liquidity

In contrast to the organized exchanges, where derivatives are standardized and prices are transparent, OTC contracts can be tailored to participants' needs and their specifications are not publicly disclosed. Although market participants can specify potential trading partners to an inter-dealer broker, they cannot precisely assess their price risk exposure. Given this lack of transparency and the risks associated with it, some energy market monitoring practices have been implemented in EU member states, but behaviors that undermine the integrity of the markets are yet to be clearly prohibited in some of the major markets.

REMIT has the objective to ensure that prices reflect a fair and competitive interplay between supply and demand and that no profits can be drawn from market abuse. According to the new rules, all market participants, including transmission system operators, suppliers, traders, producers, brokers and large users who trade in the wholesale energy markets are required to provide trading data to the Agency for the Cooperation of Energy Regulators (ACER)[1].

REMIT's obligation to publish "inside information" and prohibition of market abuse have been in force since December 2011. The Implementation Act was, however, published on 18 December 2014 and includes the reporting and registration requirements of transactions for market participants. From 7 October 2015, the obligation requires the reporting of transactions relating to the supply of electricity and natural gas with delivery in the EU, which are executed at organized marketplaces, including matched and unmatched orders. From 7 April 2016, the reporting obligation was extended to other wholesale energy market contracts (OTC standard and non-standard supply and derivative contracts, transportation contracts) and other fundamental data (e.g., planned energy generation).

Although increased transparency can provide more information on the fair price of the asset and improve the ability of regulators and practitioners to monitor and reduce transaction costs (European Commission, 2004; Boehmer et al., 2005; Bessembinder et al., 2013), the effects of improved transparency on markets

[1] ACER is the lead regulator and ultimate destination for all reports under REMIT, established by Regulation (EC) No 713/2009 of the European Parliament and Council. It is the best placed to carry out the monitoring as it has both a union-wide view of electricity and gas markets and systems in the union. National Regulatory Authorities will continue to monitor at a national level and can continue to collect additional data for national purposes.

are ambiguous (e.g. Degryse et al., 2010). For example, higher transparency may lead to lower liquidity, because better informed participants may be reluctant to post orders and give away their advantage (Harris, 1997; Madhavan et al., 2005). In addition, transparency may reduce liquidity for large transactions, but not for small transaction (Elstob, 2011). It can be argued that the publication of fundamental data removes the information advantage from being able to influence quantity traded that energy companies have had over financial investors. Furthermore, a question remains concerning the expected behavior of investors. Greater reporting implies higher administrative costs for market participants, which may increase rather than reduce transaction costs. Higher transaction costs can deter investors. A reduction of trading activity from non-physical traders could reduce liquidity, thus potentially harming the regulator goals (Council of European Energy Regulators, 2015). All in all, there are opposite views on the impact of REMIT. By analyzing the evolution of liquidity in the NBP forward market, which is the most liquid natural gas market in Europe, this study aims to identify which, if any, of these views is prevalent.

3 Method

3.1 Assembling the Database

In this study, records obtained from Tullett Prebon of transactions and quotes for the NBP forward contracts from 7 May 2010 to 29 December 2014 are considered. The database includes about a third of the total OTC market for the NBP in the period and two data sets are matched. The first includes tick-by-tick indicative quotes and the second includes tick-by-tick transaction prices and volumes. Contracts at different maturities are covered, but the focus of this study is on one-month-ahead maturity, which shows significantly higher trading frequency, and is thus meaningful when liquidity dynamics are considered (e.g. Abosedra et al., 2006).

3.1.1 Data cleaning and resampling

A stepwise cleaning procedure based on Brownlees and Gallo (2006) and Barndorff-Nielsen et al. (2009) is followed. Entries with bid, ask or execution price equal to zero, and entries with negative spreads are discarded. The trading window 7:00-17:00 (GMT) on standard working days (Monday-Friday) is considered. Simultaneous records are aggregated in a single record: transactions prices and quotes are measured by their median; volumes and number of transaction are aggregated by their respective totals. Following Barndorff-Nielsen et al. (2009), outliers are detected using a non-parametric distance-based approach: records for which the bid-ask spread is greater than 10 times the median spread of that day are discarded; records are also deleted if the midquote deviates by more than 10 mean absolute deviations from its daily-median. Thus, transaction prices are aligned to the prevailing midquote and records where the price-midquote deviation is above 10 times the median absolute deviation are discarded. Finally, the time series are adjusted to control for roll-over effects according to the calendar month delivery[2].

The data are sampled at regularly spaced time-intervals. According to Foucault et al. (2013), regular

[2]https://www.theice.com/products/Futures-Options/Energy.

time intervals are required to ensure that prices have adjusted to the information content of the cumulative transactions over time. Similarly to Zhang et al. (2005) and Boffelli and Urga (2013), the trading window is split in fixed-time intervals. For each time interval, the following information is extracted: the end-of-interval price, the end-of-interval quotes, the end-of-interval volume, the total trading volume over the interval, the total trade size over the interval, and the total number of transactions over the interval. When a time interval does not contain observations, the most recent record is used. Finally, as in Boffelli and Urga (2013), the first record of each day is excluded, because it could reflect the adjustment to the overnight information and thus display excessive variability when compared to the other observations in the day. This resampling procedure is performed at different frequencies of the midquote and transaction price series, in order to minimize volatility clustering, kurtosis and autocorrelation in the midquote and transaction price return series, which are commonly observed characteristics of high-frequency data (Engle and Russell, 2004). Accordingly, the 60-minute frequency is found to be better behaved and will be used in subsequent analysis.

3.2 Deseasonalizing and detrending market quality measures

Given seasonalities and trend that can be observed in the natural gas market (e.g. Mu, 2007), it is important to ensure that predictable market activity variation affecting the trading activity, return volatility and liquidity in a similar way is removed. In other words, the focus of the analysis should be on the irregular component (the residual series). Following, Chordia et al. (2005), the raw time series y is regressed on a set of adjustment variables, **X**, which in this study are: 11 month-of-the-year dummies (February - December); 4 day-of-the-week dummies (Tuesday-Friday); 8 hour-of-the-day dummies (8:00-15:00); a time-trend.

The residuals $\hat{u}$ from the adjustment regression $y = X\beta + u$ are standardized according the the following regression model

$$log(\hat{u}^2) = X\gamma + v. \tag{1}$$

Finally, the adjusted time series $\tilde{y}$, to be analyzed is the following

$$\tilde{y} = a + b\left(\frac{\hat{u}}{exp(X\hat{\gamma}/2)}\right), \tag{2}$$

where a and b are set so that raw and adjusted sample means and variances are the same.

3.3 Assessing associations between trading activity, return volatility and liquidity: the VAR model

Following market microstructure theory, drivers of liquidity in the one-month-ahead NBP forward market are investigated using a VAR framework that is based on the following variables:

- Order flow V_t, defined as the difference between the number of buy-initiated and sell-initiated transactions (Evans and Lyons, 2002) over the 60-minute interval from $t-1$ to t, set according to the Lee and Ready (1991)'s algorithm;

- Return volatility $|R_t|$, measured as the absolute log return from the transaction prices over the same interval;

- Liquidity, as measured by the effective half-spread $S_t = |\ln\left(\frac{P_t}{M_t}\right)\Big|$, where P_t and M_t are the transaction price and midquote recorded at the t^{th} interval, respectively (Goyenko et al., 2009). The effective spread is estimated based on the deviation of the transaction price from the midquote, where the midquote is used as proxy for the fair underlying value of the asset. Thus, this measure of liquidity can be regarded as the estimated transaction cost actually paid by a trader, as well as the gross revenue earned by the liquidity supplier (Bessembinder and Venkataraman, 2010).

Taking into account expectations that are based on microstructure theory, the VAR representation is therefore the following

$$
\begin{aligned}
V_t &= \sum_{i=1}^{p} \alpha_{V,i} V_{t-i} + \sum_{i=1}^{p} \beta_{V,i} |R|_{t-i} + \sum_{i=1}^{p} \gamma_{V,i} S_{t-i} + \varepsilon_{V,t} \\
|R|_t &= \sum_{i=0}^{p} \alpha_{|R|,i} V_{t-i} + \sum_{i=1}^{p} \beta_{|R|,i} |R|_{t-i} + \sum_{i=1}^{p} \gamma_{|R|,i} S_{t-i} + \varepsilon_{|R|,t} \\
S_t &= \sum_{i=0}^{p} \alpha_{S,i} V_{t-i} + \sum_{i=0}^{p} \beta_{S,i} |R|_{t-i} + \sum_{i=1}^{p} \gamma_{S,i} S_{t-i} + \varepsilon_{S,t},
\end{aligned} \tag{3}
$$

where the innovation terms $\varepsilon_{V,t}$, $\varepsilon_{|R|,t}$, $\varepsilon_{S,t}$ are assumed to be zero mean and mutually and serially uncorrelated at any lead and lag order. The order of lag p is selected using the Schwartz Information criterion (SIC). As, for example, in Danielsson and Payne (2012), the above model assumes contemporaneous associations between the variables. In particular, order flow is allowed to contemporaneously affect both return volatility and spread, and return volatility is allowed to influence spread. These restrictions attempt to capture the market microstructure theory based dynamics: information on the asset value is aggregated via trading activity, which subsequently affects volatility, and both volume and volatility influence liquidity (Kyle, 1985; Glosten and Milgrom, 1985; Easley and O'Hara, 1987). The imposed restrictions also ensure that the innovation terms in (3) are uncorrelated.

Given the above, the innovations in the order flow equation, $\varepsilon_{V,t}$, might reflect unpredictable changes in the demand/supply of liquidity, which can be driven by either inventory rebalancing or information-based trading. Innovations in the volatility equation, $\varepsilon_{|R|,t}$, may capture transitory, inventory-based effects, or the permanent effect of information on prices. Finally, innovations in the spread equation, $\varepsilon_{S,t}$, may mirror transitory and permanent innovations on trading activity and price volatility. The effects of these innovations are retrieved using the moving average (VMA) representation of the VAR specification, in terms of the lag operator:

$$\begin{pmatrix} V_t \\ |R|_t \\ S_t \end{pmatrix} = \begin{pmatrix} A_V(L) & B_V(L) & C_V(L) \\ A_{|R|}(L) & B_{|R|}(L) & C_{|R|}(L) \\ A_S(L) & B_S(L) & C_S(L) \end{pmatrix} \begin{pmatrix} \varepsilon_{V,t} \\ \varepsilon_{|R|,t} \\ \varepsilon_{S,t} \end{pmatrix} \tag{4}$$

where, for example, is $A_{|R|}(L) = I + A_{|R|,1}L + A_{|R|,2}L^2 + ... + A_{|R|,k}L^k$, with $L^k\varepsilon_{|R|,t} = \varepsilon_{|R|,t-k}$ defining the lag operator. Under the assumption of independence across innovation terms, the VMA representation provides the impulse responses implied by the VAR. Consequently, the lag polynomial $A_{|R|}L$ provides the cumulative effect of a unit order flow innovation on the return volatility at a k period horizon. Similarly, the cumulative effects on the spread of a unit innovation in the order flow and volatility are given by the lag polynomials A_SL and B_SL, respectively. VAR equations are estimated by OLS and inference is based on Newey-West robust standard errors. The VMA standard errors are calculated by Monte Carlo simulation, using 10,000 replications.

The above formulation assumes that the associations across variables are time-invariant. Given that these associations in the one-month-ahead NBP forward market may have changed with market conditions, a time-varying rolling approach is adopted. In particular, the VAR model in (3) is estimated assuming rolling windows of size m=4,500 (two business years) over the sample and increments between successive rolling windows of 1 period, corresponding to a 60-minute interval. This results in N=6,012 estimates of the coefficients αs, βs, γs in (3) (N=T-m+1, with T=10,521). Changes in the coefficients can therefore be assessed via plots of the rolling coefficient estimates.

3.4 Event Analysis

Since trading in the NBP OTC market is likely to have progressively adapted to meet the new reporting regulations that are effective since October 2015, the potential effect of the entering into force of REMIT in December 2011 is examined. An event analysis, inspired by Hedge and McDermott (2003), is performed to assess changes in the associations under investigation. The event at time $t = 0$ is represented by the 28^{th} of December, 2011 when REMIT entered into force[3]. In the analysis that follows, the period 7 May, 2010 to 27 December, 2011 represents the pre-REMIT time window, spanning 412 trading days in the sample. The post-REMIT time window covers the period 2 January, 2012 to 29 December, 2014, or 757 trading days in the sample. Pre- and post-REMIT estimates of the VAR model (3) are obtained over the subsamples [-413, -1] and [+1, 757], respectively. T-statistics, χ^2-statistics and the Chow test (Chow, 1960) for know structural breaks are thus used to evaluate changes in the associations between the variables after REMIT.

[3]This corresponds to the 20th day after the publication of REMIT in the Official Journal of the European Union on 8 December 2011, Regulation (EU) No. 1227/2011, Article 22, p. 15.

4 Results

4.1 Preliminary Data Analysis

As mentioned above, the analysis in this study assumes observations resampled at 60-minute frequencies, as this frequency minimizes volatility clustering, kurtosis and autocorrelation in both midquote and transaction price return series. This results in a sample of T = 10,521 observations. Descriptive statistics of order flow, volatility and spread in the one-month-ahead NBP forward market are reported in Table 1. The first four moments of the distributions are shown in columns two to five (Mean, Std.Dev., Skewness and Kurtosis, respectively); Median, first (Q_{25}) and third (Q_{75}) interquartile are shown in columns six, seven and eight. The main feature highlighted by these statistics is that, with the exception of spread, the time series have strong asymmetric distributions, while leptokurtosis appears to characterize all three series.

[Insert Table 1 here]

4.2 Deseasonalized and detrended measures

Table 2 presents the estimates from the adjustment regressions of order flow, volatility and spread in the one-month-ahead NBP forward market. Seasonal behaviors are clearly observed in the measures of market quality. Order flow is lower during the summer, reflecting the weather-dependence in the natural gas demand. Both volatility and spread are higher in the summer than in the winter. On a weekly basis, order flow appears to be higher on Tuesdays and Wednesdays; while both volatility and spread decrease within the week. Overall, there is evidence of a negative association between order flow and volatility and spread, when monthly and daily dynamics are accounted for. Conversely, when intra-day patterns are investigated, it appears that a positive association exists between these measures, such that both volatility and spread improve as order flow reduces. Furthermore, a significant negative trend is observed in all measures during the period.

[Insert Table 2 here]

The deseasonalized and detrended measures of order flow, volatility and spread are depicted in Figure 1. It appears that when predictable variations are accounted for, both volatility and spread increase in the period, mainly during 2014, when spread seems to be more volatile. This observation is supported by the descriptive statistics of the deseasonalized measures, presented in Table 3, where kurtosis (in column five) is higher when compared to the statistics of the raw series, in Table 2.

[Insert Figure 1 here]

[Insert Table 3 here]

4.3 Associations between order flow, volatility and spread in the one-month-ahead NBP forward market

Estimates of the VAR specification for the adjusted order flow, volatility and spread are given in Table 4, where estimates of the coefficients from a 9-order-lag VAR model are reported. A strong positive autocorrelation is noticed in all the three variables. Thus, there is evidence that order flow leads immediately to higher volatility, which in turn leads to significantly increased order flow. Evidence that increased volatility leads immediately to greater spread is found, whilst greater spread is associated with subsequent higher volatility. Finally, the effect of spread on subsequent order flow is mixed. These associations are apparent, not only through t-statistics for individual right-hand side(RHS) variables, but also from the χ^2-statistics in the final rows of the table, where the null hypothesis that on all coefficients of order flow, volatility and spread are simultaneously zero is tested. The adjusted R^2 coefficient in the order flow and volatility equations is approximately 0.07, whilst spread equation coefficient is approximately 0.34. Overall, results are comparable to evidence from financial markets (e.g. Chordia et al., 2005; Danielsson and Payne, 2012).

[Insert Table 4 here]

From the perspective of the VAR representation in (3), key parameters are those labelled $\alpha_{|R|}$, α_S and β_S, which capture the effect of order flow on volatility and spread, and the effect of volatility on spread, respectively. As shown in Table 4, the combined effect may be mixed as, for instance, when the cumulative effect of order flow on spread is considered. Thus, to assess the associations running across order flow, volatility and spread in the in the one-month-ahead NBP forward market, and the drivers of liquidity, the VMA representation in (4) is estimated and the cumulative impulse response functions computed.

Figure 2 shows the cumulative impulse response functions across fifty five 60-minute intervals, corresponding to five business days (along with 95% confidence intervals). The individual plots of Figure 1 represent the response of volatility and spread to a order flow shock (plots (a) and (b)) and the response of spread to a volatility shock (plot (c)). These plots suggests that the response of spread to both order flow and volatility shocks across time are greater than the immediate response. Although it appears that spread response to order flow is not significant, cumulative responses in Figure 2 imply that shocks to order flow transmit to spread across time through volatility. Hence, both volatility and spread are positively associated with order flow over time and spread is strongly positively associated with volatility. This indicates that, the higher the order flow, the higher the volatility and, in turn, the spread, the lower is liquidity in the one-month-ahead NBP forward market.

[Insert Figure 2 here]

Rolling estimates of the coefficients representing the contemporaneous associations between order flow, volatility and spread in the one-month-ahead NBP forward market ($\alpha_{|R|,0}$, $\alpha_{S,0}$, and $\beta_{S,0}$ in (3), respectively) are presented in Figure 3. The positive association between order flow and volatility appears to reduce during the period May 2010-March 2013, when volatility is low (plot(b), Figure 1). Conversely, this association increases from the second quarter to December in 2014, when volatility is higher. Thus, plot (b) of Figure 3 suggests a reduction in the association between order flow and spread during the period August 2013-December 2014, compared to the previous period in the sample. In particular, it seems that the effect of order flow on spread becomes negligible in the period 2012-14. Finally, the association between volatility and spread is positive across the sample, but shows high variability during the period 2012-14, as indicated in plot (c) of the figure. Overall, the time-varying coefficients of the VAR representation confirm that the associations between order flow, volatility and spread in the one-month-ahead NBP forward market change over time, and this may reflect variation in market conditions.

[Insert Figure 3 here]

4.4 Were associations influenced by REMIT?

Estimates of the VAR specification in the pre-REMIT and post-REMIT time periods are presented in Tables 5 and 6, respectively. Serial correlation is strong in the post-REMIT period (Table 6), in line with results from the full sample (Table 4). In the pre-REMIT period, volatility appears to be only weakly positively autocorrelated, which may indicate less uncertainty in the market. In both the subsamples, contemporaneous associations between order flow and volatility, and between volatility and spread are found, and are in line with the full sample. Nonetheless, it appears that in the post-REMIT period associations differ. In particular, a lower and less persistent association is found between spread and subsequent volatility, which is negative when compared to the pre-REMIT period. Furthermore, in contrast to the pre-REMIT and full sample, no significant effects of order flow in the spread equation are found in the post-REMIT period. These associations can be inferred from the t-statistics for the estimates of individual coefficients, as well as from the χ^2-statistics. Overall, Chow's tests on the single equations of the VAR specifications reject the null hypothesis of identical parameters across subsamples. These changes can be better assessed through the cumulative impulse responses, which are depicted in Figure 4. It is noticeable that in the post-REMIT period, the cumulative volatility response to order flow reduces. Furthermore, higher cumulative response of spread to volatility is found in the post-REMIT period, suggesting higher uncertainty in the underlying market conditions. Overall, the adjusted R^2 coefficients from the VAR specifications in the subsamples are comparable to the coefficients observed in the full sample.

[Insert Tables 5 and 6 here]

[Insert Figure 4 here]

5 Discussion

Results indicate that increases in one-month-ahead NBP order flow leads immediately to increased return volatility, which in turn leads to significantly larger spread. Hence, the findings support market microstructure theory (Glosten and Milgrom, 1985; Kyle, 1985; Easley and O'Hara, 1987) and, more broadly, the view that order flow affects asset prices. In short, the microstructure theory from financial markets can be extended to physical markets, in particular to the natural gas market, and helps in understanding the drivers of liquidity.

Price volatility in the one-month-ahead NBP forward market seems to be dependent on past trading activity, as implied by a significant cumulative response of volatility to order flow. The equilibrium volatility impulse response shows that an unexpected order flow, that is an increase in the demand of liquidity, leads to increased volatility of around 0.02% on average. Unexpected increased volatility leads to upward spread of around 0.09%. The average spread is around 0.32%, which corresponds to around 32 basis points. The estimated cumulative response of spread to volatility therefore provides a measure of the contribution of volatility in driving liquidity in the one-month-ahead NBP forward market during the period.

The magnitude of the spread in this study is comparable to what was observed by Goyenko et al. (2009) in the stock markets and by Marshall et al. (2012) in commodity markets. As argued by Payne (2003), order processing costs and inventory control costs would be greater in intermediated inter-dealer broker markets, as in this study, than in exchanges. One may therefore conjecture that spread in the one-month-ahead NBP forward market would be more likely generated by inventory costs rather than asymmetric information. Order flow would then be mainly driven by inventory rebalancing rather than the exploitation of private information. This would explain the negligible association between order flow and spread, which was observed in the rolling window estimates for the period 2012-14. Findings are therefore of interest for the costs of hedging and investment decisions, when inventory decisions and storage value are accounted for, since natural gas is a storable commodity. As argued by Felix et al. (2013), storage operators anticipate market liquidity and take it into account in their operating decisions: the lower the liquidity (the higher spread), the higher the market price, the lower is the storage value. Finally, VAR estimates suggest that the response of liquidity to price shocks may exacerbate and perpetuate price volatility, thus generating what Danielsson and Payne (2012) defined "a vicious liquidity/volatility cycle" (p. 802), and consequently are in line with evidence from financial markets.

The question whether the observed associations might be affected by the entering into force of REMIT, on the 28^{th} of December 2011, has been addressed through an event analysis. Findings from this analysis indicate the independence of spread from order flow in the post-REMIT period, also highlighted in the time-varying analysis. When the association between volatility and spread is addressed, in spite of a not statistically different spread in the subsamples, impulse response functions indicate that the equilibrium spread response to volatility is 0.07% in the pre-REMIT period, compared to 0.10% in the post-REMIT period. The implication of this finding is that although liquidity does not significantly change after REMIT, it becomes more exposed to unexpected price volatilities. Higher administrative costs, implied by REMIT, may have encouraged the gradual exit of financial investors from the commodities markets, as observed since 2013, thus contributing to price volatility in the market and increasing liquidity exposure to volatility.

Finally, it is noteworthy that the above findings are based on deseasonalized and detrended series, and thus the predictable changes in the natural gas market activity affecting liquidity and volatility in similar ways is excluded. Conclusions are based on a share of the one-month-ahead OTC market in the period, and future studies may consider a longer period and markets at different maturities. In addition, results may be sensible to reorderings of the three variables in the VAR representation. Furthermore, the database does not discriminate between commercial and financial investors, and therefore an exhaustive assessment of the impact of REMIT on the quality of the NBP/European natural gas forward market cannot be made. In this respect, access to the data that will become available to regulators via REMIT would significantly help in understanding energy markets and their drivers.

6 Conclusions

This study examined the associations between order flow, volatility and spread, and was aimed at identifying the drivers of liquidity in the one-month ahead NBP forward market. It relied on expectations based on the market microstructure theory and previous literature in finance. By allowing for associations to be time-varying in response to some unexpected changes in the underlying market conditions, it accounted for the fact that natural gas markets in Europe are relatively young and are evolving. Moreover, the rules that regulate the market are also evolving, and therefore the effect of the entering into force of REMIT on the one-month-ahead market quality was also investigated.

Results indicated a positive association between spread and volatility, that is a negative association between liquidity and volatility in the market. In particular, the equilibrium response of spread to unexpected price volatility changes was estimated and found to be related to the inventory control problem. Drivers of liquidity in the one-month-ahead NBP forward market were identified, and there are indications that factors influencing price volatility changes will impact liquidity. These observations are of interest to the energy sector and market players, who consider liquidity as an effective way to spread correct price signals about the fundamental values of demand and supply. Findings are also valuable for independent regulators and energy policy makers, who are interested in monitoring the efficiency and competitiveness of the market.

No significant changes in the associations between the measures of market quality considered in this study was observed following the entering into force of REMIT. However, there was evidence of greater exposure of liquidity to unexpected price changes after REMIT, which might reflect fewer investors in the market. Overall, drawing from a large share of transactions in the one-month-ahead forward NBP market, the present study identified potential drivers of liquidity in the European gas markets. The findings supported the extension of the market microstructure theory to the physical markets, and thus contributed towards understanding liquidity dynamics in energy markets, and their underlying driving forces.

References

Abosedra, S., Elkhal, K., and Al-Khateeb, F. (2006). Forecasting performance of natural gas futures market: An assessment of recent data. *Journal of Business and Economics Research*, 4:65–70.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time series effects. *Journal of Financial Markets*, 5:31–56.

Amihud, Y. and Mendelson, H. (1980). Dealership market: Market-making with inventory. *Journal of Financial Economics*, 8:31–53.

Amihud, Y. and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17:223–249.

Bagehot, W. p. (1971). The only game in town. *Financial Analysts Journal*, 27:12–14.

Banti, C., Phylaktis, K., and Sarno, L. (2012). Global liquidity risk in the foreign exchange market. *Journal of International Money and Finance*, 31:267–291.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N. (2009). Realised kernels in practice: Trades and quotes. *Econometrics Journal*, 12:1–33.

Benston, G. and Hagerman, R. (1974). Determinants of bid-asked spreads in the over-the-counter market. *Journal of Financial Economics*, 1:353–364.

Bessembinder, H. (1994). Bid-ask spreads in the interbank foreign exchange markets. *Journal of Financial Economics*, 35:317–348.

Bessembinder, H., Maxwell, W. F., and Venkataraman, K. (2013). Trading activity and transaction costs in structured credit products. *Financial Analysts Journal*, 69:55–67.

Bessembinder, H. and Venkataraman, K. (2010). *Bid-Ask spreads: Measuring trade execution costs in financial markets*. In: Encyclopedia of Quantitative Finance, edited by Rama Cont, John Wiley & Sons, 2010.

Boehmer, E., Saar, G., and Yu, L. (2005). Lifting the veil: an analysis of pre-trade transparency at the NYSE. *Journal of Financial Markets*, 8:217–264.

Boffelli, S. and Urga, G. (2013). Macroannouncements, bond auctions and rating actions in the european government bond spreads. *Centre of Econometric Analysis, Cass Business School*, Working Paper XX.

Brennan, M., Chordia, T., and Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, 49:345–373.

Brennan, M. and Subrahmanyam, A. (1996). Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. *Journal of Financial Economics*, 41:441–464.

Brownlees, C. and Gallo, G. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics and Data Analysis*, 51:2232–2245.

Chordia, T., Roll, R., and Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics*, 56:3–28.

Chordia, T., Roll, R., and Subrahmanyam, A. (2002). Order imbalance, liquidity, and market returns. *Journal of Finance*, 65:111–130.

Chordia, T., Sarkar, A., and Subrahmanyam, A. (2005). The joint dynamics of liquidity, returns, and volatility across small and large firms. *Staff Report, Federal Reserve Bank of New York, No. 207, available at http://www.econstor.eu/.*, pages 111–130.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28:591–604.

Council of European Energy Regulators (2015). Annual Report 2015. AR2015.

Cummins, M. and Murphy, B. (2015). *Natural gas markets and products*. In: Handbook of Multi-Commodity Markets and Products, Andrea Roncoroni, Gianluca Fusai, Mark Cummins (Eds.), John Wiley & Sons Ldt, 135-180.

Danielsson, J. and Payne, R. (2012). Liquidity determination in an order driven market. *The European Journal of Finance*, 18:799–821.

Degryse, H., deJong, F., and Van Kervel, V. (2010). The Impact of MiFID on the Quality of Euronext. *Working paper, Tilburg University*.

Easley, D. and O'Hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19:69–90.

Elstob, P. (2011). FSA at odds with European Commission over aspects of MiFID II. WBC. 29. March 2011, available at http://commoditymkts.wordpress.com/.

Engle, R. and Russell, J. (2004). *Analysis of high frequency data*. In: Ait-Sahalia and L. P. Hansen (Eds.): Handbook of Financial Econometrics, Vol 1: Tools and Techniques, 383-426. 2010, Elsevier Science: North-Holland.

European Commission (2004). Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004, on Markets in Financial Instruments Amending Council Directives 85/611/EC and 93/6/EEC and Directive 200/12/ EC of the European Parliament and of the Council and repealing Council Directive 93/22/EEC. Official Journal of the European Union 145, 30/4/2004.

European Commission (2011). Proposal for a Directive of the European Parliament and of the Council on Markets in Financial Instruments Repealing Directive 2004/39/EC of the European Parliament and of the Council. , Brussels, COM (2011) 656 final, Oct. 2011.

European Commission (2015). Quarterly Report on European Gas Markets. Volume 7 (issue 4; fourth quarter of 2014).

Evans, M. and Lyons, R. (2002). Order flow and exchange rate dynamics. *Journal of Political Economy*, 110:170–180.

Felix, B., Woll, O., and Weber, C. (2013). Gas storage valuation under limited market liquidity: an application in Germany. *European Journal of Finance*, 19:715–733.

Foucault, T., Pagano, M., and Röell, A. (2013). *Market liquidity*. Oxford University Press, 2013.

Garman, M. (1976). Market microstructure. *Journal of Financial Economics*, 3:257–275.

Glosten, L. and Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71–100.

Goyenko, R., Holden, C., and Trzcinka, C. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics*, 92:153–181.

Harris, L. (1997). Order Exposur eand Parasitic Traders. *Working paper, Marshall School ofBusiness*, 1:1–22.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance*, 46:179–207.

Hasbrouck, J. (2009). Trading costs and returns for US equities: estimating effective costs from daily data. *Journal of Finance*, 64:1445–1477.

Hedge, S. and McDermott, J. (2003). The liquidity effects of revision to the S&P index: An empirical analysis. *Journal of Financial Markets*, 6:413–459.

International Gas Union (2015). *Wholesale Gas Price Survey*. International Gas Union, May 2015.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, 53:1315–1335.

Larosière, J. (2009). Final report of the high-level group on financial super-vision in the EU, February 2009.

Lee, C. and Ready, M. (1991). Inferring trade direction from intraday data. *Journal of Finance*, 46:733–746.

Madhavan, A., Porter, D., and Weaver, D. (2005). Should securities markets be transparent? *Journal of Financial Markets*, 8:266–288.

Marshall, B., Nguyen, N., and Visaltanachoti, N. (2012). Commodity liquidity measurement and transaction costs. *Review of Financial Studies*, 25:599–638.

Mu, X. (2007). Weather, storage, and natural gas price dynamics: fundamentals and volatility. *Energy Economics*, 29:46–63.

Nijman, L. (2012). The impact of new wave of financial regulation for European energy markets. *Energy Policy*, 47:468–477.

O'Hara, M. (1995). *Market microstructure theory*. Cambridge, MA: Basil Blackwell.

Pastor, L. and Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 113:642–685.

Payne, R. (2003). Informed trade in spot foreign exchange markets: an empirical investigation. *Journal of International Economics*, 61:307–329.

Pilipovic, D. (2007). *Energy risk: Valuing and managing energy derivatives*. McGrawHill Publishers (2nd ed.), NewYork.

Stoll, H. R. (1978). The supply of dealer services in security markets. *Journal of Finance*, 33:1133–1151.

Zhang, L., Mykland, P., and Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100:1394–1411.

Table 1: Descriptive statistics of order flow, volatility and spread in the one-month-ahead NBP forward market

	Mean	Std. Dev.	Skewness	Kurtosis	Median	Q_{25}	Q_{75}
Order flow	6.855	7.927	2.401	15.829	4.000	1.000	10.000
Volatility	0.311	0.548	2.361	265.859	0.170	0.000	0.398
Spread	0.003	0.004	0.000	9.507	0.002	0.001	0.004

Note: The table reports the descriptive statistics of the order flow, volatility and spread, resampled at 60-minute frequency, corresponding to a sample T of 10,521 observations.Order flow is defined as the difference between the number of buy-initiated and sell-initiated transactions over the 60-minute interval from $t-1$ to t; Volatility is measured as the absolute log return from the transaction prices over the same interval; Spread is the effective half-spread, which is based on the deviation of the transaction price from the midquote, used a proxy for the true underlying value of the asset, and can be regarded as an estimate of the transaction cost actually paid by a dealer, as well as the gross revenue earned by the liquidity supplier. The first four moments of the distributions are shown in columns two to five (Mean, Std.Dev., Skewness and Kurtosis, respectively); Median, first (Q_{25}) and third (Q_{75}) interquartile are shown in columns six, seven and eight, respectively.

Table 2: Estimates from the adjustment regressions of order flow, volatility and spread in the one-month-ahead NBP forward market

		Order flow			Volatility			Spread		
Month		Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat
	February	0.635	0.657	0.966	0.051***	0.021	2.385	0.0004	0.0003	1.337
	March	-1.454***	0.614	-2.368	0.039*	0.021	1.853	0.0004	0.0003	1.229
	April	-0.660	0.605	-1.090	0.057**	0.026	2.225	0.0013***	0.0003	4.423
	May	-1.708***	0.550	-3.103	0.046**	0.021	2.153	0.0020***	0.0004	4.846
	June	-2.526***	0.570	-4.433	0.050**	0.023	1.914	0.0019***	0.0004	4.391
	July	-2.579***	0.540	-4.776	0.045*	0.025	1.801	0.0017***	0.0004	4.073
	August	-2.342***	0.560	-4.181	0.052***	0.022	2.423	0.0018***	0.0003	5.420
	September	-1.367***	0.564	-2.425	0.081***	0.033	2.450	0.0012***	0.0003	3.704
	October	-0.542	0.585	-0.926	0.020	0.025	0.798	0.0009***	0.0003	2.989
	November	-1.096*	0.598	-1.832	-0.012	0.018	-0.667	0.0003	0.0004	0.819
	December	-4.616***	0.498	-9.268	-0.016	0.020	-0.823	0.0004	0.0003	1.196
Day of the week										
	Tuesday	1.121***	0.274	4.095	-0.045**	0.020	-2.270	-0.0006***	0.0002	-3.875
	Wednesday	0.783***	0.314	2.497	-0.066***	0.020	-3.262	-0.0008***	0.0002	-4.498
	Thursday	0.182	0.302	0.604	-0.089***	0.019	-4.683	-0.0008***	0.0002	-5.142
	Friday	0.049	0.285	0.172	-0.070***	0.020	-3.426	-0.0007***	0.0002	-4.379
Hour of the day										
	8.00	-0.667**	0.335	-1.994	0.296***	0.024	12.33	0.0029***	0.0002	15.562
	9.00	-0.054	0.367	-0.146	0.067***	0.023	2.878	0.0001	0.0001	1.067
	10.00	-2.665***	0.313	-8.512	-0.0001	0.024	-0.006	-0.0003***	0.0001	-2.370
	11.00	-3.833***	0.317	-12.089	-0.084***	0.015	-5.772	-0.0006***	0.0001	-5.149
	12.00	-5.593***	0.296	-18.907	-0.128***	0.016	-8.272	-0.0007***	0.0001	-6.281
	13.00	-5.382***	0.271	-19.829	-0.101***	0.013	-7.726	-0.0006***	0.0001	-5.644
	14.00	-3.736***	0.296	-12.614	-0.054***	0.013	-4.095	-0.0004***	0.0001	-3.404
	15.00	-1.730***	0.323	-5.360	-0.012	0.016	-0.760	-0.0002***	0.0001	-2.339
Trend		-0.00032***	0.00003	-9.08166	-0.00002***	0.000002	-8.005	-0.000035***	0.000003	-11.100

Note: The table reports the coefficients from the adjustment regressions of the order flow, volatility and spread on 11 month-of-the-year dummies (February - December); 4 day-of-the-week dummies (Tuesday-Friday); 8 hour-of-the-day dummies (8:00-15:00); a time-trend. Robust standard errors are based on Newey-West estimator. ***, **, * denote 1%, 5% and 10% significance level, respectively.

Table 3: Descriptive statistics of the deseasonalized and detrended order flow, volatility and spread measures

	Mean	Std. Dev.	Skewness	Kurtosis	Median	Q_{25}	Q_{75}
Order flow	6.855	7.927	2.574	18.900	4.722	1.400	9.911
Volatility	0.311	0.548	3.811	353.133	0.180	0.029	0.432
Spread	0.003	0.004	0.000	30.824	0.002	0.001	0.004

Note: The table reports the descriptive statistics of the adjusted order flow, volatility and spread. The first four moments of the distributions are shown in columns two to five (Mean, Std.Dev., Skewness and Kurtosis, respectively); Median, first (Q_{25}) and third (Q_{75}) interquartile are shown in columns six, seven and eight, respectively.

Table 4: Estimates of the VAR specification for order flow, volatility and spread

Equation	Order flow			Volatility			Spread		
Regressor	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat
Intercept	0.4524***	0.027	16.62	0.2338***	0.038	6.15	0.1424***	0.051	2.80
V_t	-	-	-	0.2121***	0.012	17.60	0.0123	0.012	1.03
V_{t-1}	0.1603***	0.013	3.40	-0.0516***	0.008	-6.14	-0.0132	0.012	-1.10
V_{t-2}	0.0486***	0.015	2.70	-0.0044	0.007	-0.60	0.0035	0.009	0.37
V_{t-3}	0.0520***	0.011	0.03	-0.0198***	0.007	-2.69	0.0078	0.013	0.59
V_{t-4}	-0.0093	0.010	2.23	-0.0056	0.008	-0.72	-0.0169**	0.009	-1.93
V_{t-5}	0.0209**	0.008	0.13	-0.0113	0.007	-1.54	-0.0035	0.008	-0.43
V_{t-6}	0.0085	0.010	2.17	-0.0049	0.007	-0.68	-0.0066	0.008	-0.79
V_{t-7}	0.0399***	0.008	-0.09	-0.0237***	0.007	-3.38	0.0009	0.009	0.09
V_{t-8}	0.0539***	0.009	-1.11	-0.0092	0.007	-1.29	0.0214**	0.009	2.46
V_{t-9}	0.0725***	0.010	0.27	-0.0131	0.008	-1.55	-0.0058	0.010	-0.60
$\lvert R_t\rvert$	-	-	-	-	-	-	0.2050***	0.039	5.28
$\lvert R_{t-1}\rvert$	0.0455***	0.010	4.52	0.0331***	0.011	3.12	-0.0162	0.013	-1.26
$\lvert R_{t-2}\rvert$	0.0411***	0.010	4.08	0.0433***	0.013	3.23	-0.0114	0.012	-0.93
$\lvert R_{t-3}\rvert$	0.0003	0.010	0.03	0.0257***	0.010	2.48	-0.0226**	0.011	-2.04
$\lvert R_{t-4}\rvert$	0.0223**	0.010	2.21	0.0351***	0.012	3.01	-0.0044	0.009	-0.51
$\lvert R_{t-5}\rvert$	0.0010	0.010	0.10	0.0638***	0.012	5.12	0.0003	0.011	0.03
$\lvert R_{t-6}\rvert\rvert$	0.0221**	0.010	2.19	0.0513***	0.013	3.82	0.0072	0.010	0.71
$\lvert R_{t-7}\rvert$	-0.0007	0.010	-0.07	0.0317***	0.010	3.23	-0.0099	0.010	-0.97
$\lvert R_{t-8}\rvert$	-0.0097	0.010	-0.96	0.0321***	0.011	2.99	-0.0051	0.009	-0.59
$\lvert R_{t-9}\rvert$	0.0027	0.010	0.27	0.0357***	0.010	3.70	-0.0052	0.009	-0.56
S_t	-	-	-	-	-	-	-	-	-
S_{t-1}	0.0022	0.013	0.17	0.0088	0.011	0.78	0.4032***	0.029	13.7
S_{t-2}	-0.0005	0.015	-0.04	0.0177	0.012	1.52	0.0449**	0.019	2.42
S_{t-3}	-0.0169	0.013	-1.29	0.0025	0.011	0.23	0.0495***	0.011	4.37
S_{t-4}	-0.0056	0.013	-0.44	0.0238**	0.012	2.03	0.0211*	0.012	1.70
S_{t-5}	-0.0467***	0.011	-4.08	0.0001	0.011	0.01	0.0550***	0.016	3.49
S_{t-6}	0.0007	0.012	0.05	-0.0058	0.012	-0.49	0.0196	0.014	1.39
S_{t-7}	-0.0020	0.010	-0.19	0.0059	0.013	0.47	0.0452***	0.015	3.08
S_{t-8}	-0.0066	0.011	-0.58	0.0106	0.010	1.030	0.0084	0.014	0.60
S_{t-9}	0.0228**	0.012	1.92	0.0237**	0.011	2.089	0.0936***	0.017	5.40
Adjusted R^2			0.070			0.073			0.344
χ^2 Order flow			395.9***			322.4***			14.48
χ^2 Volatility			43.25***			65.94***			34.41***
χ^2 Spread			38.78***			18.29**			304.1***

Note: The table reports coefficients from a 9-lag VAR involving order flow, volatility, defined as the absolute returns and spreads over the full sample, corresponding to T=10,512 observations. Order flow V_t, defined as the difference between the number of buy-initiated and sell-initiated transactions over the 60-minute interval from $t-1$ to t; Volatility $|R_t|$, measured as the absolute log return from the transaction prices over the same interval; Spread is identified by the effective half-spread and represents the deviation of the transaction price from the midquote, used a proxy for the true underlying value of the asset, and can be regarded as an estimate of the transaction cost actually paid by a dealer, as well as the gross revenue earned by the liquidity supplier. The series upon which the VAR is estimated are resampled on at 60-minute frequencies. All t-statistics and χ^2-statistics are based on Newey-West heteroscedasticity and autocorrelation robust standard errors. χ^2-statistics in the final rows of the table test the null hypothesis that coefficients on all included order flow, volatility or spread variables are simultaneously zero. ***, ** and * denote significance at 1%, 5% and 10%, respectively.

Table 5: Estimates of the VAR specification for order flow, volatility and spread: Pre-REMIT time period

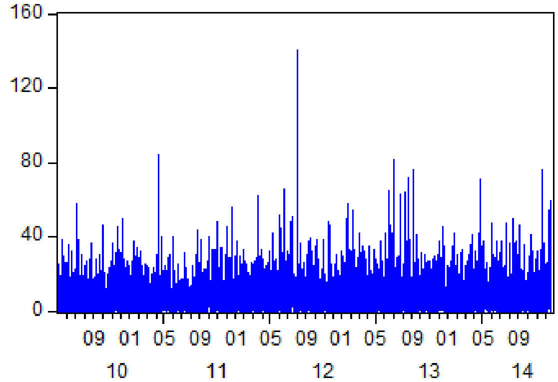
Equation	Order flow			Volatility			Spread		
Regressor	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat
Intercept	0.4082***	0.043	9.54	0.2143***	0.056	3.82	0.2109***	0.029	7.40
V_t	-	-	-	0.2564***	0.016	15.6	0.0269	0.019	1.39
V_{t-1}	0.1836***	0.020	9.09	-0.0591***	0.015	-3.95	-0.0275*	0.015	-1.80
V_{t-2}	0.0293	0.019	1.58	0.0013	0.015	0.09	0.0100	0.013	0.75
V_{t-3}	0.0385**	0.019	2.05	-0.029**	0.012	-2.39	-0.0161	0.013	-1.20
V_{t-4}	-0.0273	0.020	-1.39	0.0036	0.014	0.26	-0.0197	0.013	-1.52
V_{t-5}	0.0354**	0.018	1.98	0.0196	0.014	1.36	-0.0002	0.013	-0.02
V_{t-6}	-0.0141	0.017	-0.84	-0.0057	0.014	-0.42	-0.0133	0.013	-0.99
V_{t-7}	0.0463**	0.019	2.40	-0.0142	0.014	-0.99	-0.0245*	0.013	-1.89
V_{t-8}	0.0602***	0.019	3.19	0.0140	0.014	1.03	0.0286**	0.013	2.21
V_{t-9}	0.0615***	0.018	3.34	-0.0214	0.014	-1.59	-0.0051	0.013	-0.39
$\|R_t\|$	-	-	-	-	-	-	0.2093***	0.054	3.89
$\|R_{t-1}\|$	0.0509***	0.013	3.80	0.0272	0.023	1.18	0.0017	0.023	0.07
$\|R_{t-2}\|$	0.0427***	0.018	2.36	0.0235	0.023	1.01	-0.0295**	0.014	-2.11
$\|R_{t-3}\|$	0.0133	0.015	0.88	0.0157	0.017	0.95	-0.0156	0.011	-1.39
$\|R_{t-4}\|$	0.0334**	0.016	2.13	0.0224	0.015	1.53	0.0019	0.012	0.15
$\|R_{t-5}\|$	0.0153	0.014	1.11	0.0273*	0.014	1.92	-0.0006	0.014	-0.04
$\|R_{t-6}\|$	0.0579***	0.022	2.59	0.0314*	0.013	2.50	0.0176	0.021	0.85
$\|R_{t-7}\|$	0.0064	0.011	0.58	0.0174	0.013	1.34	0.0014	0.021	0.07
$\|R_{t-8}\|$	0.0210	0.018	1.17	0.0220	0.016	1.39	0.0016	0.014	0.11
$\|R_{t-9}\|$	0.0436**	0.017	2.57	0.0020	0.013	0.16	-0.0017	0.011	-0.15
S_t	-	-	-	-	-	-	-	-	-
S_{t-1}	-0.0020	0.023	-0.09	0.0047	0.024	0.19	0.2962***	0.032	9.33
S_{t-2}	-0.0294	0.023	-1.27	0.0633***	0.024	2.68	0.0921***	0.025	3.68
S_{t-3}	-0.0302	0.024	-1.27	0.0069	0.021	0.33	0.0410**	0.021	1.95
S_{t-4}	0.0138	0.025	0.54	0.0281	0.020	1.44	0.0470***	0.018	2.63
S_{t-5}	-0.0435*	0.022	-1.95	-0.0216	0.021	-1.03	0.0437**	0.024	1.84
S_{t-6}	0.0022	0.023	0.09	0.0684***	0.024	2.81	0.0239	0.023	1.02
S_{t-7}	-0.0183	0.021	-0.86	-0.0201	0.021	-0.94	0.0248	0.021	1.19
S_{t-8}	-0.0179	0.022	-0.82	0.0073	0.020	0.36	-0.0100	0.019	-0.53
S_{t-9}	0.0083	0.024	0.35	0.0323*	0.018	1.82	0.0721***	0.018	4.00
Adjusted R^2			0.089			0.072			0.271
χ^2 Order flow			190.4***			287.1***			16.52*
χ^2 Volatility			41.36***			13.31			29.82***
χ^2 Spread			17.43**			20.82**			314.2***

Note: The table reports coefficients from a 9-lag VAR involving order flow, volatility, defined as the absolute returns and spreads, estimated in the pre-REMIT time window May 2010-December 2011, corresponding to T=3,708 observations. Order flow V_t, defined as the difference between the number of buy-initiated and sell-initiated transactions over the 60-minute interval from $t-1$ to t; Volatility $|R_t|$, measured as the absolute log return from the transaction prices over the same interval; Spread is identified by the effective half-spread and represents the deviation of the transaction price from the midquote, used a proxy for the true underlying value of the asset, and can be regarded as an estimate of the transaction cost actually paid by a dealer, as well as the gross revenue earned by the liquidity supplier. The series upon which the VAR is estimated are resampled on at 60-minute frequencies. All t-statistics and χ^2-statistics are based on Newey-West heteroscedasticity and autocorrelation robust standard errors. χ^2-statistics in the final rows of the table test the null hypothesis that coefficients on all included order flow, volatility or spread variables are simultaneously zero. ***, ** and * denote significance at 1%, 5% and 10%, respectively.

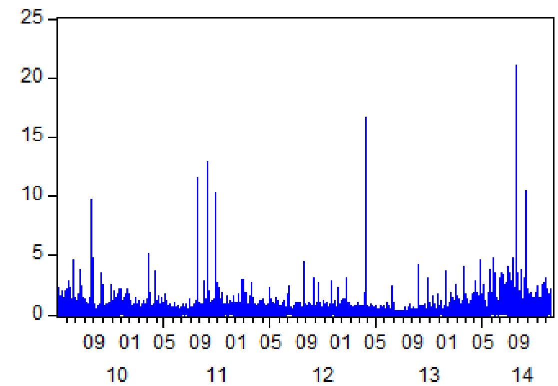
Table 6: Estimates of the VAR specification for order flow, volatility and spread: Post-REMIT time period

Equation	Order flow			Volatility			Spread		
Regressor	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat	Coeff	Std.Er.	t-Stat
Intercept	0.4959***	0.0354	14.02	0.2265***	0.0477	4.75	0.1312**	0.0652	2.01
V_t	-	-	-	0.1953***	0.0148	13.23	0.0050	0.0150	0.34
V_{t-1}	0.1466***	0.0195	7.51	-0.0467***	0.0100	-4.69	-0.0079	0.0159	-0.49
V_{t-2}	0.0542***	0.0137	3.96	-0.0055	0.0084	-0.66	0.0001	0.0121	0.01
V_{t-3}	0.0557***	0.0145	3.85	-0.0135	0.0092	-1.47	0.0175	0.0178	0.98
V_{t-4}	-0.0038	0.0124	-0.30	-0.0064	0.0095	-0.68	-0.0176	0.0112	-1.57
V_{t-5}	0.0140	0.0128	1.10	-0.0223***	0.0083	-2.68	-0.0067	0.0100	-0.67
V_{t-6}	0.0146	0.0131	1.11	-0.0047	0.0087	-0.54	-0.0039	0.0104	-0.37
V_{t-7}	0.0358***	0.0131	2.73	-0.0273***	0.0080	-3.41	0.0113	0.0124	0.91
V_{t-8}	0.0509***	0.0144	3.54	-0.0192**	0.0085	-2.25	0.0169	0.0111	1.53
V_{t-9}	0.0761***	0.0147	5.16	-0.0075	0.0107	-0.71	-0.0056	0.0126	-0.44
$\|R_t\|$	-	-	-	-	-	-	0.2017***	0.0520	3.88
$\|R_{t-1}\|$	0.0461***	0.0197	2.34	0.0340***	0.0108	3.14	-0.0182*	0.0139	-1.31
$\|R_{t-2}\|$	0.0443**	0.0204	2.17	0.0486***	0.0163	2.99	0.00005	0.0163	0.003
$\|R_{t-3}\|$	-0.0026	0.0140	-0.18	0.0271**	0.0133	2.04	-0.0258**	0.0160	-1.61
$\|R_{t-4}\|$	0.0185	0.0129	1.44	0.0379**	0.0165	2.30	-0.0069	0.0114	-0.61
$\|R_{t-5}\|$	-0.0051	0.0091	-0.56	0.0807***	0.0156	5.18	0.0031	0.0158	0.20
$\|R_{t-6}\|$	0.0056	0.0091	0.62	0.0543***	0.0190	2.86	0.0034	0.0112	0.31
$\|R_{t-7}\|$	-0.0014	0.0102	-0.14	0.0361***	0.0138	2.62	-0.0116	0.0112	-1.04
$\|R_{t-8}\|$	-0.0226**	0.0091	-2.48	0.0342**	0.0139	2.46	-0.0055	0.0106	-0.52
$\|R_{t-9}\|$	-0.0161	0.0102	-1.58	0.0492***	0.0138	3.57	-0.0033	0.0128	-0.26
S_t	-	-	-	-	-	-			
S_{t-1}	0.0003	0.0158	0.02	0.0144	0.0129	1.12	0.4364***	0.0347	12.60
S_{t-2}	0.0072	0.0179	0.40	0.0029	0.0133	0.22	0.0230	0.0226	1.02
S_{t-3}	-0.0139	0.0158	-0.88	0.0044	0.0128	0.35	0.0557***	0.0138	4.03
S_{t-4}	-0.0113	0.0148	-0.76	0.0206	0.0141	1.46	0.0097	0.0157	0.62
S_{t-5}	-0.0491***	0.0133	-3.69	0.0118	0.0138	0.86	0.0607***	0.0190	3.19
S_{t-6}	-0.0009	0.0143	-0.06	-0.032**	0.0130	-2.49	0.0136	0.0168	0.81
S_{t-7}	0.0028	0.0119	0.24	0.0195	0.0150	1.30	0.0522***	0.0179	2.92
S_{t-8}	-0.0057	0.0135	-0.42	0.0100	0.0118	0.84	0.0105	0.0170	0.62
S_{t-9}	0.0257*	0.0137	1.87	0.0234*	0.0136	1.72	0.0967***	0.0212	4.57
Adjusted R^2			0.062			0.075			0.365
χ^2 Order flow			228.3***			182.5***			8.741
χ^2 Volatility			27.57***			55.52***			18.16**
χ^2 Spread			32.66***			20.79**			273.7***

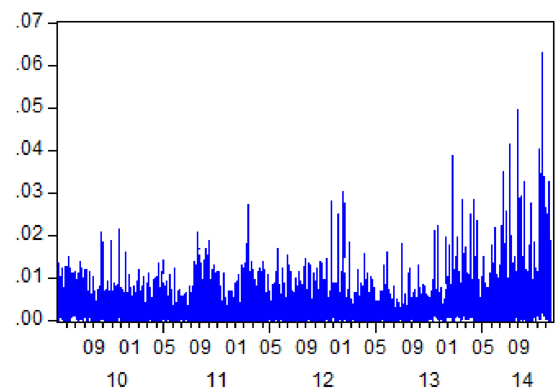
Note: The table reports coefficients from a 9-lag VAR involving order flow, volatility, defined as the absolute returns and spreads, estimated in the post-REMIT time window January 2012-December 2014, corresponding to T=6,804 observations. Order flow V_t, defined as the difference between the number of buy-initiated and sell-initiated transactions over the 60-minute interval from $t-1$ to t; Volatility $|R_t|$, measured as the absolute log return from the transaction prices over the same interval; Spread is identified by the effective half-spread and represents the deviation of the transaction price from the midquote, used a proxy for the true underlying value of the asset, and can be regarded as an estimate of the transaction cost actually paid by a dealer, as well as the gross revenue earned by the liquidity supplier. The series upon which the VAR is estimated are resampled on at 60-minute frequencies. All t-statistics and χ^2-statistics are based on Newey-West heteroscedasticity and autocorrelation robust standard errors. χ^2-statistics in the final rows of the table test the null hypothesis that coefficients on all included order flow, volatility or spread variables are simultaneously zero. ***, ** and * denote significance at 1%, 5% and 10%, respectively.



(a) Order flow



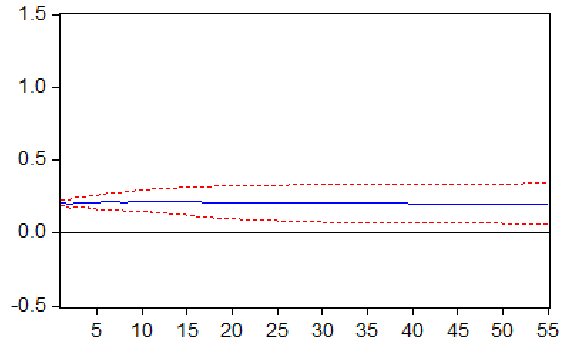
(b) Volatility



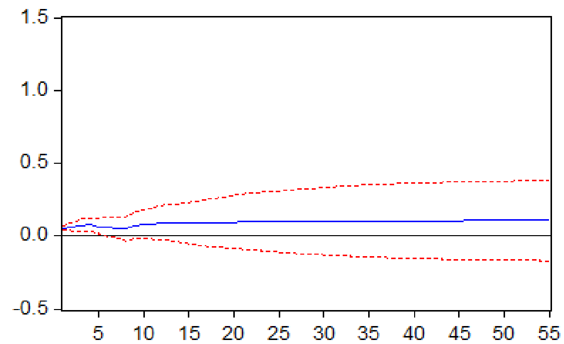
(c) Spread

Figure 1: Deseasonalized and detrended measures of order flow, volatility and spread

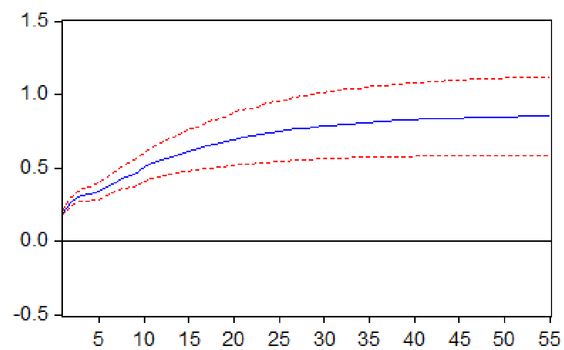
Order flow is defined as the difference between the number of buy-initiated and sell-initiated transactions over the 60-minute interval from $t-1$ to t; Volatility is measured as the absolute log return from the transaction prices over the same interval; Spread is the effective half-spread, which is based on the deviation of the transaction price from the midquote, used a proxy for the true underlying value of the asset, and can be regarded as an estimate of the transaction cost actually paid by a dealer, as well as the gross revenue earned by the liquidity supplier.



(a) Volatility to order flow



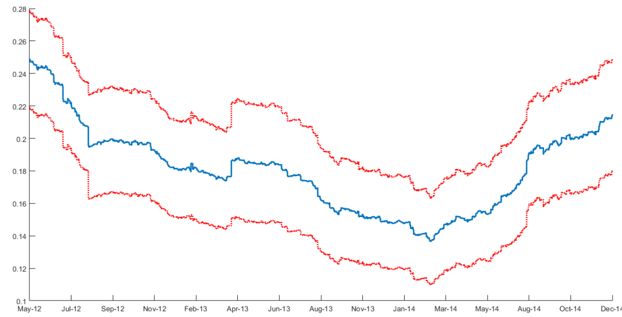
(b) Spread to order flow



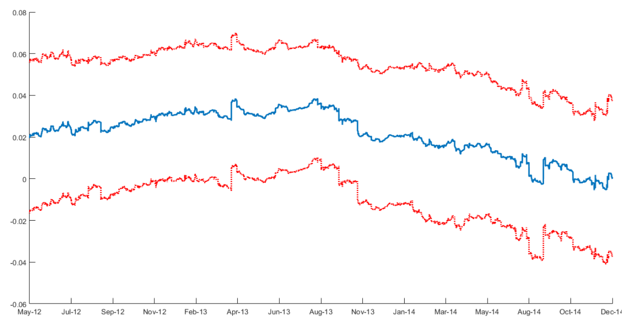
(c) Spread to volatility

Figure 2: Cumulative impulse response functions

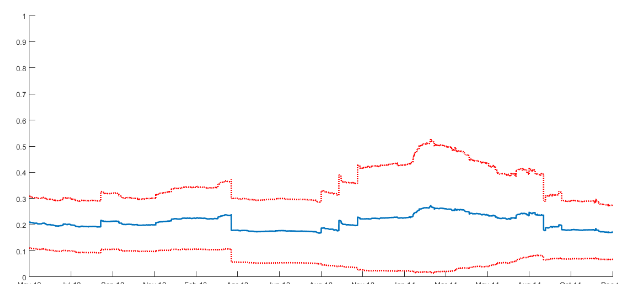
The figure shows: (a) Cumulative volatility response to a unit order flow shock; (b) Cumulative spread response to a unit order flow shock; (c) Cumulative spread response to a unit volatility shock. The impulse response functions are calculated from the VAR representation in Eq. (4). The blue solid line in each panel gives the actual impulse response function and the red dotted lines are the 95% confidence interval estimated using Monte Carlo simulations of the VAR model with 10,000 replications. The x-axis values give the number of 60-minute intervals since the shock was felt.



(a) Association between order flow and volatility



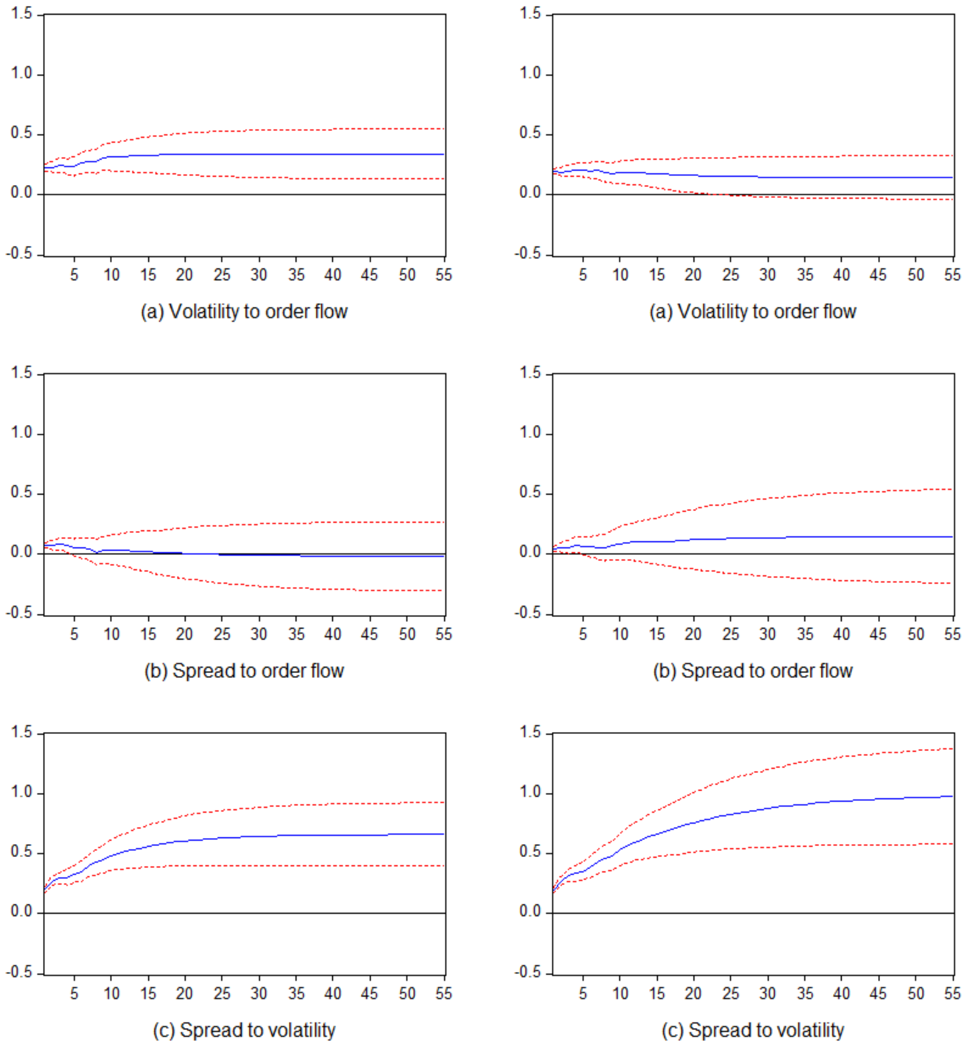
(b) Association between order flow and spread



(c) Association between volatility and spread

Figure 3: Time-varying associations between market quality measures in the one-month ahead NBP forward market

The figure shows time-varying representing the contemporaneous associations between: (a) Order flow and volatility; (b) Order flow and spread; (c) Volatility and spread ($\alpha_{|R|,t}$, $\alpha_{S,t}$, and $\beta_{S,t}$ in Eq. (3), respectively). The blue solid line in each panel gives the rolling coefficients and the red dotted lines represents the 95% confidence interval.



Pre-REMIT Post-REMIT

Figure 4: Cumulative impulse response functions in the pre- and post-REMIT time periods

The figure shows: (a) Cumulative volatility response to a unit order flow shock; (b) Cumulative spread response to a unit order flow shock; (c) Cumulative spread response to a unit volatility shock in the pre- and post-REMIT time periods. The impulse response functions are calculated from the VAR representation in Eq. (4). The blue solid line in each panel gives the actual impulse response function and the red dotted lines are the 95% confidence interval estimated using Monte Carlo simulations of the VAR model with 10,000 replications. The x-axis values give the number of 60-minute intervals since the shock was felt.