Clients' Connections

Measuring the Role of Private Information in Decentralised Markets*

Péter Kondor
London School of Economics
CEPR

Gábor Pintér
Bank of England

July 16, 2019

Abstract

We propose a new measure of private information in decentralised markets – connections – defined as the number of dealers with whom a client trades in a time period. Using proprietary data for the UK government bond market, we show that clients have systematically better performance when having more connections, and this effect is stronger during macroeconomic announcements. Time-variation in market-wide connections also helps explain yield dynamics. Given our novel measure, we present two applications suggesting that (i) dealers pass on information, acquired from their informed clients, to their subsidiaries, and (ii) informed clients better predict the order-flow intermediated by their dealers.

JEL Classification: G12, G14, G24

Keywords: Private Information, Client-Dealer Connections, Government Bond Market

*First version: May 2018. Previous version was titled "Private Information and Client Connection in Government Bond Markets". We would like to thank Andy Blake, Marco Di Maggio, Wouter Den Haan, Dong Lou, Daniel Paravisini, Angelo Ranaldo, Lucio Sarno, Eric Stafford, Gertjan Vlieghe and participants at the LSE seminar, the Conference on Non-Bank financial institutions and the 2019 SED Conference (St. Louis) for helpful comments. Special thanks to Lena Boneva for helping us understand the ZEN data, and to Marcin Kacperczyk (discussant) for many fruitful conversations. The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. Kondor acknowledges financial support from the European Research Council (Starting Grant \#336585). Email: p.kondor@lse.ac.uk and gabor.pinter@bankofengland.co.uk.

1 Introduction

A main role of financial markets is to aggregate private information held by economic agents. Trading activity and subsequent adjustments in asset prices release this information to the wider public, thereby making markets more efficient and increasing the welfare of society. The main challenge facing any scientific study of this mechanism is that neither private information nor the identities of its owners are readily observable.

Our paper proposes a proxy for private information. We combine a detailed dataset of the UK government bond market, covering the identities and transactions of trading parties, with insights from the microstructure literature. The idea is that, just like in centralised markets where informed traders may split their trades over time to slow down information revelation and avoid market-impact (Kyle, 1985), informed traders in decentralised markets may submit orders to different dealers at the same time, thereby splitting their trades in the cross-section. This implies that one should observe a trader obtaining private information to trade with more dealers than usual. Accordingly, our proposed proxy for private information in decentralised markets is the time-variation in the number of dealers that clients trade with, which we will refer to as the clients' *connections*.[1]

Our empirical analysis yields two sets of results. First, we confirm that connections serve as a proxy for private information by showing that (i) clients make more profitable trades when having more connections and (ii) time-variation of total client connections in the market helps explain daily changes in yields. Second, we present two of the many possible applications of our proxy: (i) we find suggestive evidence that dealers learn from their informed clients and pass this information to their subsidiaries, and (ii) we also show that the nature of private information proxied by connections pertains to future order-flow, i.e. more connected clients better predict the order-flow intermediated by the dealers they trade with.

To organise ideas and to guide our empirical strategy, we start our analysis by proposing a simple model in the spirit of Glosten and Milgrom (1985). The model formalises the insight that trading with more dealers may be advantageous because it helps the client hide her private information. This, however, requires the client to reach out for quotes

[1] Our study focuses on the UK government bond market, because (i) being one of the most liquid decentralised markets, it provides a particularly hard test to measure private information, (ii) our dataset provides a detailed, almost universal coverage of all transactions on this market, and (iii) the government bond market plays a crucial role in the economy as the yield curve serves as a benchmark in many financial transactions, it affects government financing costs and plays an important role for the implementation of monetary policy.

from more dealers, which is costly. Therefore, the client will do so only when the benefit of hiding information is sufficiently large, that is, when her information is sufficiently precise. In these periods, the client should overperform.

To illustrate the viability of this idea, we first analyse trading around the Brexit referendum. Given the large uncertainty before the vote, market participants were motivated to either reduce their exposure radically, or to generate private information and bet on the outcome. In line with our hypothesis, we show that a change in their number of dealer connections helps identify the client group with private information. In particular, the group of clients who were connected with more dealers on the day before the referendum persistently increased the duration of their positions for days before the referendum and, subsequently, outperformed other clients when the yield curve dropped immediately as the outcome of the poll became public.

Then we turn to systematic evidence. We expect that when a client is connected to more dealers, her trades are more profitable even after controlling for the volume and the number of her transactions in the given period. This effect should not be driven by favourable transaction prices, but by forecasting future price movements. That is, the price of gilts that connected clients buy (sell) should increase (decrease) in subsequent days. We also expect this effect to be more pronounced in more information sensitive periods, e.g. around macroeconomic announcements. We find empirical evidence for each of these predictions. Including client and time fixed effects, we identify these results from the time-series of a given client's activity.

Moreover, we consider aggregate implications for yield dynamics. We construct a market-wide measure of private information – the total number of client-dealer connections in the system in a time period. We then measure the response of yields to changes in aggregate connections, and find a significant effect even after controlling for trading volume and the total number of clients in the market.

Given our proxy for private information, we offer two of the many possible applications. As a first application, we provide suggestive evidence that dealers pass on information, acquired from their informed clients, to their subsidiaries. To show this, we use a novel source of variation in our data: for each dealer, we are able to distinguish between trading accounts that perform a market-making function from trading accounts that correspond to other, client-like arms of the given dealer bank, i.e. the given dealer's *subsidiaries*. We then test whether dealers' subsidiaries perform better when the given dealer trades with a larger proportion of high-connection clients. We find that this is indeed the case, suggesting that these subsidiaries obtain the information that their dealers learn from informed clients.

As a second application, we search for the source of private information captured by the time-variation in clients' connectedness. We find that more connected dealers can better predict the maturity structure of other clients' order-flow, especially the part of the order-flow received by their own dealers in subsequent days. For instance, when a more connected dealer orders is concentrated on the short-end of the yield curve in a given day, her dealer is more likely to receive a disproportionate share of orders for short bonds in the following five days. We also show that trading in line with the maturity structure of clients' future orders can be profitable because of the resulting pressure on prices. A limitation of this analysis is that we cannot observe whether the client is gathering information from the quotes she receives from her dealers, or dealers leak the information to its best clients.

Finally, we present further extensions and additional robustness checks that include (i) re-estimating our empirical model on data at daily (instead of monthly) frequency, (ii) considering alternative performance measures and additional decompositions, (iii) changing the definition of macroeconomic announcements, (iv) considering other market-wide measures of asymmetric information that aggregate connections explain, (v) revisiting the role of the centrality of dealers.

Related Literature Our paper is the first to suggest clients' connections as a measure of private information in decentralized markets. Our study is related to several streams of the literature.

There is a vast literature on measuring private information in financial markets. A large group of these papers focus on security based measures (e.g. Easley, Kiefer, O'Hara, and Paperman, 1996; Chakravarty, Gulen, and Mayhew, 2004; Duarte and Young, 2009; Roll, Schwartz, and Subrahmanyam, 2010; Johnson and So, 2018). These papers identify securities for which a large share of transactions are likely to be motivated by private information in a given period, typically using the aggregate volume characteristics of those securities, and study the implied return patterns. Instead, our measure allows to study informed transactions of any given client. As our applications show, this feature changes the range of relevant questions we can address with our approach.

A more related group of papers identify informed transactions focusing on the activity of a specific group of clients such as large shareholder activists or corporate insiders (Cohen, Malloy, and Pomorski, 2012; Collin-Dufresne and Fos, 2015) often during specific episodes (Boulatov, Hendershott, and Livdan, 2013; Hendershott, Livdan, and Schurhoff, 2015). By design, these studies are mostly focusing on the cross-sectional heterogeneity in information, building on ex-ante assumptions of which clients should be more informed and in which periods private information should be concentrated.[2] Instead, we use time-series heterogeneity to identify client specific periods of informed trading. That is, our measure can systematically identify periods of informed trading for any given client, even if these periods are uncorrelated across clients.

Our first application studies potential information leakages across clients, their dealers, and the subsidiaries of these dealers.[3] While there are many empirical works studying the trading process in decentralised markets (e.g. Gabrieli and Georg, 2014; Hollifield, Neklyudov, and Spatt, 2017; Li and Schurhoff, 2019) most of these do not focus on the role of private information.[4] Instead, the most related work to this application is Maggio, Franzoni, Kermani, and Sommavilla (2019). Just as we do in this application, they use the network of transactions across market participants to study the flow of private information among them. Apart from the context – they focus on brokers and their clients in stock markets – their proxy of informed trades and the suggested mechanism are also different from our approach. They identify a client's informed transactions as those which are executed by a more connected broker. The argument is that central brokers gather information by executing informed trades, which is then leaked to their best clients through these transactions. Instead, we identify informed transactions as those which are executed when the client is more connected. Our argument is that the client chooses to be more connected when her information is more precise in order to hide it.[5]

Our second application is related to the literature on price discovery in government bond markets (Fleming and Remolona, 1999; Balduzzi, Elton, and Green, 2001; Green, 2004; Brandt and Kavajecz, 2004; Pasquariello and Vega, 2007; Hortacsu and Kastl, 2012; Valseth, 2013). This literature emphasises the informational role of clients' and/or dealers' orderflow. We add to this literature by highlighting the empirical link between variation in connections and order flow predictability. We are able to do so due to the

[2]Another approach is to study the effect of transactions which ex-post turns out to be private information driven. For instance (Meulbroek, 1992; Kacperczyk and Pagnotta, 2019) investigates the effect of transactions that subsequently became subject to SEC investigations of insider trading activities.

[3]There is a related, growing theoretical literature on the role of private information in decentralized markets such as Duffie, Malamud, and Manso (2009), Golosov, Lorenzoni, and Tsyvinski (2014), Babus and Kondor (2018), Brancaccio, Li, and Schurhoff (2017) amongst others.

[4]A notable exception is Hagstromer and Menkveld (2019) which uses short-term comovement across quotes of different dealers to map information percolation.

[5]To make this difference salient, we show in Section 5.1 that the results in our first application are robust to the inclusion of dealers' centrality as a control. Also, in Section 6.6 we show that the variation in clients' connections and in their dealers' centrality is largely independent.

important feature of our dataset: for each trade we can observe the identity of both parties. This allows us to map out the dynamics of connections of government market participants and explore their links with the price discovery process.

The remainder of the paper is as follows: Section 2 introduces the environment, concepts and hypotheses illustrated by the example of financial betting around the Brexit referendum; Section 3 describes the data sources and provides summary statistics; Section 4 presents the empirical results on using connections as proxy for private information; Section 5 presents the two applications of our measure; Section 6 presents robustness checks and further extensions; Section 7 concludes.

2 Concepts and Hypotheses

We start this section with a basic description of the micro-structure of the UK gilt market. Then, with the illustration of a simple model, we discuss our main hypotheses.

2.1 Primary Dealers in the UK Gilt Market

The key actors in the UK gilt market are the primary dealers, also known as gilt-edged market markers (GEMMs). In our sample period between 2011 and 2017, their number fluctuates between 20 and 24. From now on, we refer to this group as dealers. The UK Debt Management Office (DMO) tenders new issues of government securities to dealers. Clients, as asset managers, commercial banks and foreign central banks buy and sell government securities mostly through bilateral transactions to this group.[6] Primary dealers are committed to make, on demand, continuous and effective two-way prices to their clients by regulation. They must also maintain a minimum market share (DMO, 2011).[7]

When a client trades in the UK gilt market, she can observe quotes of all dealers on electronic trading platforms. However, these observed quotes are merely indicative and only small trades can be executed at these prices. If the client wishes to trade a larger quantity, she directly contacts the dealers typically via the phone. Unlike other, centralised exchanges (e.g. the UK gilt futures market) that are increasingly automated, the gilt cash market, which our study focuses on, continues to retain its traditional OTC characteristics where reputation and trading relationships matter largely for dealers (to

[6] In our sample, only about 1% of client trades are directly between clients.

[7] See Benos and Zikes (2018) for further details about the institutional arrangements of the UK gilt market.

continue to attract order flow and thereby trading profitably) as well as for clients (to receive favourable price quotes).

In our sample, we observe that clients tend to trade with a relatively small and persistent subset of all the dealers. In practice, this subset corresponds closely to the subset they requests quotes from. Based on interviews with traders, we understand that clients perceive that asking quotes from many dealers can be costly.[8] In particular, the main (perceived or real) cost of asking for quotes, but not trading with some of the dealers is that it might damage the relationship between the client and the given dealers. For example, a dealer might feel that she gives out information on her inventory when providing tight quotes. This information might be used against the dealer. If this is not reciprocated with executed trades, the dealer might decide to give less informative, that is, less tight quotes to that particular client next time.

2.2 The Mechanism and Main Implications

Our main conjecture is that the time-variation in clients' connections can be a proxy for the time-varation in the precision of their private information. The underlying mechanism is that, just like in centralised markets where informed traders may split their trades over time to slow down information revelation and avoid market-impact (Kyle, 1985), informed traders in decentralised markets may submit orders to various dealers, thereby splitting their trades in the cross-section. However, this requires the client to reach out for quotes from those dealers, which might be costly. The client will do so only when the advantage is large, that is, when its information is sufficiently precise. This implies that one should observe a trader obtaining more precise private information to trade with more dealers than usual.[9] This mechanism provides a number of testable predictions.

First, consider the time-variation of the performance of a given client. Under our main conjecture, we should observe that when clients are connected to more dealers, they overperform. However, overperformance could come from multiple sources. For instance,

[8] Moreover, even the dealer whose quote is accepted by the client pays some informational cost, as all the other dealers who have also been requested to provide quotes will know that the transaction took place. (In fact, the runner-up in the auction gets informed specifically that her quote was the second best.) Especially in the case of a large transaction, the dealers whose quotes were not accepted might use this price and quantity information against the dealer (with the accepted quote) when she tries to manage the resulting change in her inventory in the inter-dealer market.

[9] Since Kyle (1985), the micro structure literature has extensively studied how private information can be concealed by splitting informed orders in smaller amounts over time to avoid market impact (e.g. Garleanu and Pedersen, 2013; Mascio, Lines, and Naik, 2017; Back, Collin-Dufresne, Fos, Li, and Ljungqvist, 2018)

even if connections were not related to information, clients requesting more quotes would confront their dealers to more competition, possibly resulting in more favourable transaction prices. Instead, according to our conjectured mechanism, we should expect that the price of government bonds, purchased by the client in these periods, should increase in subsequent days compared to the price of bonds they sell. That is, a more connected client's overperformance should come from the correlation of the direction of their transactions and future price movements. Finally, we expect these differential effects to be more pronounced in more information-sensitive periods, for example, around important macroeconomic announcements. The idea is that in these periods informed traders signals are more precise. We summarize these predictions in the following hypotheses.

Hypothesis 1 *More connections for a client i in a given interval should be associated with higher trading profit.*

Hypothesis 2 *This relationship would be stronger in periods with more precise private information, i.e. around public announcements and news events.*

Hypothesis 3 *More connections for a client i in a given interval should be associated with a stronger connection between her buy (sell) transactions and subsequent positive (negative) returns.*

Second, we consider implications to the price discovery process. In the absence of news, innovations in the yield curve should be driven by private information. Also, under our conjecture, average connections in a given time period is a measure of the amount of private information present in the market. Therefore, we should expect a comovement between this measure and innovations in the yield curve. This gives our last hypothesis.

Hypothesis 4 *Periods with higher aggregate connections should be associated with larger absolute innovations in yields.*

While these predictions are intuitive, it is important to show that they are consistent with a rational framework. For instance, in equilibrium dealers might foresee that an unusual request for a quote from a client might imply that that client has private information. Their provided quote should adjust to this belief. For this purpose, we build a simple model of trading and network formation in Appendix A. The model is a variant of Glosten and Milgrom (1985). Informed clients and uninformed liquidity traders interacting with market makers. The new element of the trading protocol is that clients can decide whether to seek bid and ask quotes from one or more risk neutral, competitive market makers in each round. Sampling quotes from more market makers is costly. After observing the quotes, clients can decide which dealer to trade with. The informativeness of clients' signals varies in the time-series and in the cross-section. We assume that announcements correspond to periods with more informative signals for many clients.

In equilibrium, each client requests quotes only from her regular dealer when her information precision is low and from multiple dealers when it is high. In the latter case, she receives identical quotes and she mixes her choice of a dealer. As a result, in periods with higher information precision she will trade with more dealers in expectation. We carefully derive the results supporting each hypothesis above.

In Section 4, we test these hypotheses and find empirical evidence for each one of them in the data.

3 Measurement and Summary Statistics

In this section we describe the data and construct the two main variable of interest: clients' connections and performance.

3.1 Data Source

To analyse how the dynamics of client-dealer connections are related to clients' trading performance and information, one needs a detailed transaction-level dataset which contains information on the identity of both sides of a trade. The proprietary ZEN database maintained by the UK Financial Conduct Authority (FCA), fittingly provides this information together with information on the transaction date and time; the execution price and quantity; the International Securities Identification Number (ISIN); the account number, the buyer-seller flag. The ZEN database contains trade reports for all secondary-market transactions, where at least one of the counterparties is an FCA-regulated entity. We focus exclusively on conventional gilts. Given that all dealers in our sample are FCA-regulated, we have at least one report for each dealer-client transaction, thereby giving us virtually full coverage of the client trade universe. Our sample covers the period between October 2011 and June 2017. We match our transaction-level data with information on bond duration and end-of-day closing prices obtained from Datastream.

A key aspect of our empirical analysis is to exploit the time-variation in client-dealer connections, which requires the matching of each transaction with a client identifier. The names of clients are recorded as unstructured strings of text in the ZEN database. Moreover, a typical client tends to have multiple accounts with different client names across accounts and also within the same account. We use a textual algorithm that searches the unstructured strings of names and accounts, and assigns a unique client identifier to each transaction. When constructing client identifiers, we aim at the highest possible level of consolidation by treating parent companies, subsidiaries and different arms as one client. We end up with 474 identified clients and about 1.67 million trades transacted by them. The trading activity of these clients covers around 80% of all client activity (in terms of trading volume) in the UK gilt market.

3.2 Client-Dealer Connections

Our baseline measure of connections is the number of dealers a given client is connected to in a given time period. A client is connected to a dealer if she trades with the dealer at least once.[10] Since client connectivity is a key variable in our analysis, we provide some descriptive statistics to describe it.

Table 2 presents summary statistics based on our baseline regression sample that is aggregated to the client-month level. We find that the average client in a given month is connected to four dealers and carries out about 19 transactions with them. There is substantial sample variation: the average difference in connections between the 90th and 10th percentile is 11. To illustrate how much of the variation in client connectivity is a cross-sectional phenomenon, we compute the averages of our measures at the client-level, and plot the resulting distribution in a histogram (Top Row of Figure 5). We find that the distribution of the connectivity measure is positively skewed, with the mass of clients having low values and a few clients exhibiting large values.

Clients that are on average more connected can differ from less connected clients along other time-invariant characteristics such as size, business model etc. To control for this, we purge out client fixed effects from our connectivity measures and plot the resulting distribution in a histogram (Bottom Row of Figure 5). We find substantial within-client variation: the average difference in connections between the 90th and 10th percentile is 4.5, which is non-negligible compared to the corresponding value using across-client

[10] To check for the robustness of our results, we also use eigenvector centrality (Bonacich and Lloyd, 2001) as an alternative measure of connectivity. This measure, used in recent papers (Maggio, Kermani, and Song, 2017), not only takes into account the number of dealers a given client trades with but also the number of other clients that are connected to those dealers that the given client trades with.

variation (7.5). Similarly, the standard deviation of first-order connections is around 3.3 in the cross-section and still as high as 1.9 when using only the within-client variation. This substantial *within-variation* in connections is an important feature of the data, which our empirical analysis will primarily rely on.

3.3 Trading Performance

3.3.1 Baseline Measure

To measure trading performance, we follow Maggio, Franzoni, Kermani, and Sommavilla (2019) and compute the T-day-horizon return on each trade of client i in month t, measured as the percentage difference between the transaction price and the closing price T days after the transaction date.[11] Formally, for each trade j, we construct the measure $Performance_{i,t}^{T}$ as follows:

$$Performance_j^T = \left[\ln\left(P^T\right) - \ln\left(P_j^{\star}\right)\right] \times \mathbf{1}_{B,S}, \tag{3.1}$$

where $P_j^{\star}$ is the transaction price, P^T is the T-day ahead closing price of the corresponding gilt, and $\mathbf{1}_{B,S}$ is an indicator function equal to 1 when the transaction is a buy trade, and -1 when it is a sell trade. All transactions-specific returns are then averaged within month t using the pound value of the trades as weights. As robustness, we also present the results using unweighted monthly average returns.

Table 2 summary statistics of the 3-day and 5-day (weighted and unweighted) performance measures. Panel B shows that average performance is significantly larger for clients with more dealer connections compared to clients with fewer connections. More importantly, as shown in Panel C, we also find that the average client performs significantly better in months with more dealer connections compared to months when the same client has fewer connections. For example, the average client has a 1.5bp higher 5-day performance in high-connection months compared to low-connection months.

3.3.2 Decomposing Trading Performance

We now propose a decomposition method which extends our baseline performance measurement. The T-day performance of a client on a trade can be high because the given

[11]The T-day horizon starts at the start of each day and ends after T days. We use overlapping time windows. For example, to compute one-day performance measures ($T = 1$), we compare all trades on day 1 to the closing price on day 2, and compare all trades on day 2 to the close price on day 3, and so on.

client faces lower price impact compared to other clients trading at the same time. We refer to this as the transaction component of performance. Alternatively, trading performance can be high because the given client can better anticipate future prices changes. We refer to this as the anticipation component of performance. Building on 3.1, we compute the decomposition for each transaction j as follows:

$$\ln\left(P^T\right) - \ln\left(P_j^\star\right) \approx \underbrace{\left[\ln\left(P^T\right) - \ln\left(\overline{P}\right)\right]}_{Anticipation} + \underbrace{\left[\ln\left(\overline{P}\right) - \ln\left(P_j^\star\right)\right]}_{Transaction}, \tag{3.2}$$

where $\overline{P}$ is the only new term which denotes the average transaction price (based on all available dealer-client trades in the corresponding gilt) measured around the time of transaction j. To estimate $\overline{P}$, we experiment with two time definitions. First, we use all relevant trades on the day of transaction j to compute the average transaction price $\overline{P}$. As a second, more accurate measure, we split each trading day into three parts, and compare the transaction price to the corresponding one of the three intra-day averages.[12] Given the trade-level decomposition, we then collapse our dataset at the client-month level using both volume-weighted and unweighted monthly average returns.

Note that most of the recent empirical work on financial networks (Afonso, Kovner, and Schoar, 2014; Hendershott, Li, Livdan, and Norman, 2017; Hollifield, Neklyudov, and Spatt, 2017; Maggio, Kermani, and Song, 2017) focused on the transaction component. Distinguishing between the transaction component and the anticipation component allows us to test whether more connections increase performance because clients can achieve more favourable deals (at lower mark-ups) or because clients have private information about future price changes.

4 Client Connections as Proxy for Private Information

This section presents our main empirical results, supporting the key message of our paper: time-variation in client-dealer connections can be used to proxy time-variation in private information. We proceed in three steps. First, we illustrate the viability of the hypotheses laid out in Section 2.2 by taking a closer look at the gilt trading activity around the Brexit referendum. Second, we turn to clients' performance and supporting evidence for

[12]The intra-day time windows are <11am, 11am-15pm and >15pm, which are set to have an approximately even number of transactions across the time windows.

Hypotheses 1–3. Using panel-data regressions, we show that clients' trading performance systematically increases when the given client trades with more dealers. Third, we study innovations in yields and Hypothesis 4. We provide evidence that variation in total client-dealer connections in the market comove with the day-to-day innovations in the level and slope of the yield curve.

4.1 Betting on Brexit: An Event Study

As a motivating example, we take a close look at the connectedness-performance relationship during the days around the Brexit referendum on leaving the European Union. The referendum took place on Thursday 23 June 2016, and the results that 51.9% of the participants voted to leave became public on Friday morning (24 June 2016). Based on polls, the chances of a leave or a remain vote were close to 50-50 leaning slightly towards remain immediately before the vote. Either way market prices were expected to jump. In particular, the common perception was that a leave result would likely trigger a rate cut soon, leading to an immediate downward shift in the yield curve on 24 June. Indeed, this is what happened with the 1-year, 5-year and 25-year yields dropping by 14bp, 30bp and 24bp, respectively, on 24 June. This was followed by the Bank of England cutting the policy rate by 25bp in August.

Given the large uncertainty before the vote, market participants were motivated to either reduce their exposure radically, or to generate private information and bet on the outcome. Reportedly, major hedge funds ordered private opinion polls to generate an informational edge.[13] Our main hypothesis implies that we should be able to separate these two groups from each other based on the change of their connectivity before the vote. We should see that clients with private information increase the number of dealers they trade with to hide this information. Furthermore, they should be the group who, in average, increases the duration of its portfolio to speculate on the leave outcome and when the yield curve eventually drops, they should overperform the others.

[13] Reportedly, major hedge funds ordered private opinion polls to generate an informational edge for this bet and earned handsomely on those bets:

> "Behind the scenes, a small group of people had a secret – and billions of dollars were at stake. Hedge funds aiming to win big from trades that day had hired YouGov and at least five other polling companies [...]. Their services, on the day and in the days leading up to the vote, varied, but pollsters sold hedge funds critical, advance information, including data that would have been illegal for them to give the public. Some hedge funds gained confidence, through private exit polls, that most Britons had voted to leave the EU, or that the vote was far closer than the public believed – knowledge pollsters provided while voting was still underway and hours ahead of official tallies." ("The Brexit Short: How Hedge Funds Used Private Polls to Make Millions", Bloomberg Businessweek, 25th June, 2018)

To verify this hypothesis, we group all those private clients who traded on the referendum day 23 June into two groups based on the following client-specific measure:

$$\alpha_i = connections_{i,Jun23} - \overline{connections}_i, \tag{4.1}$$

where $connections_{i,Jun23}$ is the number of dealers that client i traded with on the day of the referendum; the term $\overline{connections}_i$ is the average daily connectivity of client i based on the whole sample (2011 Oct – 2017 Jun). The variable α_i captures whether the given client, on the referendum day, had unusually high or low connectivity compared to its own long-run average. We identify 126 private clients who traded on the day

Table 1: Summary Statistics of the 126 Clients Trading on 23 June 2016

Client Type	Number of Clients	α Mean	Volume Mean	Net Duration Mean	5-day Performance Mean	5-day Performance Median
Low-α	63	-0.80	14.1m	5.5m	-0.425	-0.013
High-α	63	0.96	25.2m	107m	0.434	0.327

Notes: this table provides descriptive statistics of the 126 identified private clients that traded on 23 June 2016. These clients are placed in two groups depending on whether their α is below (top row) or above (bottom row) the median value of α. Performance is measured in 100*log points. The variable α is measured in terms of number of connections. Volume and Net Duration are in £.

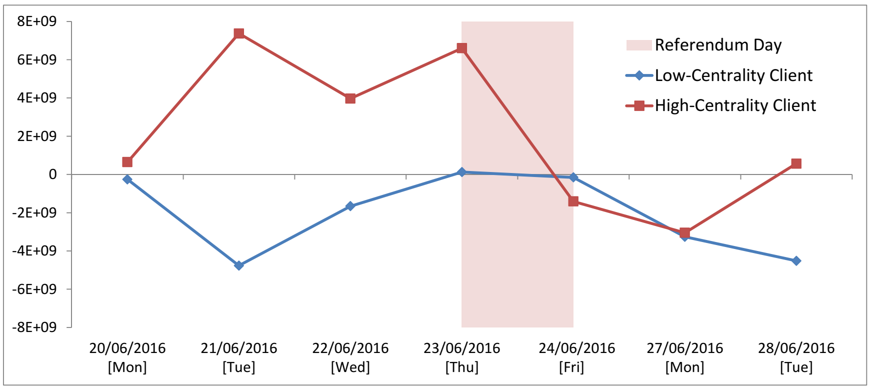
of the referendum, and Table 1 provides summary statistics of their performance and connectedness. High-α clients traded with approximately one (0.96) additional dealer compared to their respective average. In turn, low-α clients traded with approximately one (0.8) fewer dealer compared to their respective average.

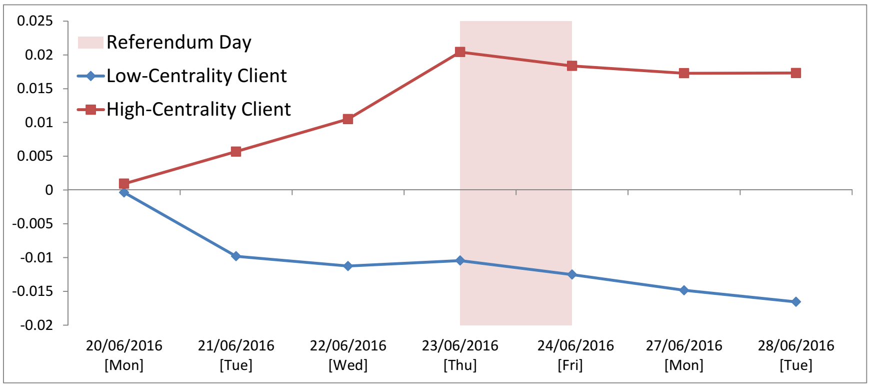
We find that high-α clients performed much better on the referendum day compared to low-α clients. For example, the mean 5-day performance of the high-α clients was about 43bp which was more than 86bp higher than the mean performance of the low-α clients (-43bp). This is primarily due to the fact that high-α clients substantially increased their long position in gilts before prices increased sharply in the following days: the average high-α client's change in net duration was about 20 times (107m) that of the low-α client (5.5m), with this difference being much sharper than the difference in trading volume across the two groups.

We illustrate these stylised facts in panels (a) and (b) of Figure 1, showing the duration and cumulative performance of the two groups in the days around the vote. While this episode is intuitive, note that the differences in performance of the high- and low-α groups might come from other, unobserved heterogeneity in these two groups. Indeed, it is quite likely that the traders who decide to bet on the outcome of the Brexit vote are very

Figure 1: Connectivity and Performance around the Brexit Referendum

(a) Aggregate Daily Net Duration of High-α and Low-α Clients





(b) Cumulative Returns of Low-α and High-α Clients.

Notes: In Panel a, the red squared line depicts the evolution of the duration-weighted net position of those 63 clients that have high within-connectedness (high α 4.1) on the day of the referendum. The blue circled line evolution of the duration-weighted net position of those 63 clients that have low within-connectedness (low α 4.1) on the day of the referendum. In Panel b, the red squared line depicts the cumulative average returns of those 63 clients that have high within-connectedness (high α 4.1) on the day of the referendum. The blue circled line depicts cumulative average returns of those 63 clients that have low within-connectedness (low α 4.1) on the day of the referendum. The average returns for both groups are weighted by the individual clients' daily trading volume. The returns are computed using the closing price on 29 June 2016 as the reference price.

different from those institutions that decide to cut back their exposure in this volatile period. Also, this particular episode might be special. Hence we turn to systematic evidence in the next Section, where we can include client- and time- fixed effects as well as additional controls to decompose the different forces at play.

4.2 Client Profitability

In this part, we connect the time-variation in clients' connections with the time-variation in their performance along the lines of Hypotheses 1–3.

4.2.1 Baseline Results

To estimate whether a client's trading performance increases when the client increases its connections with the primary dealer sector, we run the following monthly panel regression:

$$Performance_{i,t}^{T} = \beta \times ClientConnections_{i,t} + X_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}, \tag{4.2}$$

where $Performance_{i,t}^{T}$ is the trading performance (3.1) of client i in month t at horizon T; $ClientConnections_{i,t}$ is the number of dealers the given client is connected to in month t; α_i and μ_t are client and time fixed effects; $X_{i,t}$ includes controls such as the number of transactions and trading volume. These controls are important for checking that our connections variable is not simply picking up the effect of increased trade size (Easley and O'Hara, 1987; Merrick, Naik, and Yadav, 2005; Maggio, Franzoni, Kermani, and Sommavilla, 2019).

Throughout the analysis, in computing standard errors we take the most conservative approach, and employ two-way clustering at the client and time level. This allows for arbitrary correlation across time and across clients. Our baseline results use data at monthly frequency. This is to reduce measurement noise and also to avoid oversampling those clients who trade actively, possibly on most trading days. Nevertheless, as will be shown in Subsection 6.1, our results are robust to using data at daily frequency as well.

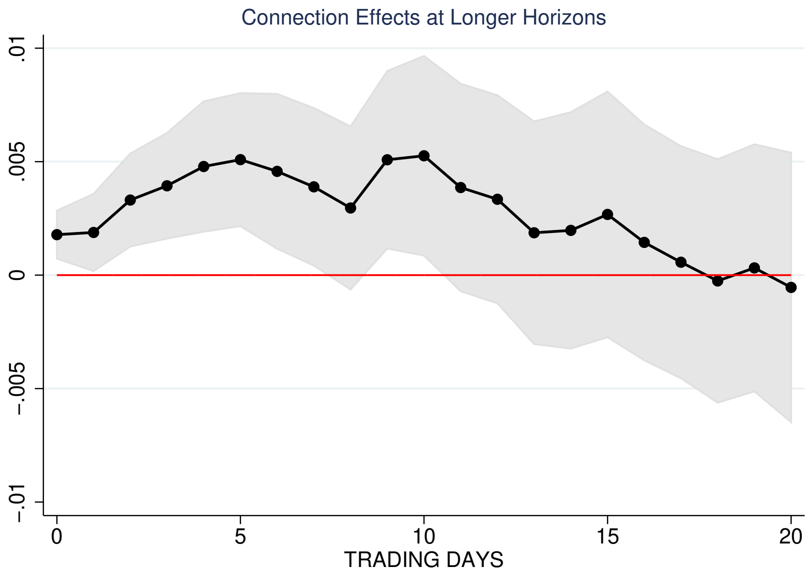
The main coefficient of interest in 4.2 is β which captures the relation between client connections and trading performance. Table 3 reports our baseline results with panel A and panel B showing the results for value-weighted and unweighted trading performance, respectively. Each column corresponds to a different trading horizon going from $T = 0$ to $T = 5$. We find a positive relationship between client connections and trading performance, which is statistically significant at every horizon for both types of performance measures. Moreover, we find little evidence that variation in a client's trading volume or number of transactions would affect the given client's trading performance.[14]

The results are also economically significant. For example, using the estimate (0.48bp) in Column 6 of Table 3, we find that if a client's connections increase in a given month

[14]Figure 19 in the Appendix shows the results from a pooled regression with client fixed effects excluded. While client connections continue to have a significantly positive relationship with trading performance, the coefficients on trading volume and transaction number also appear statistically significant in the cross-section. This may be explained by the fact that our clients include a range of investor types (e.g. insurance companies, hedge funds, pension funds). Also, the cross-sectional distribution of size and transaction number may be correlated with other characteristics, as studied extensively by the mutual fund literature (Elton, 1993; Chen, Jegadeesh, and Wermers, 2000; Kacperczyk, Sialm, and Zheng, 2005).

by one, then her trading performance doubles relative to her mean (we are using the fact that the median and mean 5-day returns are 0.45bp and 0.43bp, respectively). Table 4 further illustrates the economic significance of the performance-connection relationship. Panel 4a compares months when clients have low connectedness to months when they have high connectedness. Single-sorting using the within-variation in connections, we find that the difference in median performances is about 0.5bp, consistent with our baseline regression results (Table 3). Moreover, clients trade much more when they are more connected: the median trading volume is about £15million (£53million) in months when the client has fewer (more) dealer connections than its sample average. The performance difference coupled with the difference in trading volume in high and low connectivity months implies that the majority of positive trading performance is concentrated in high connectivity months. To reinforce that our results are not simply picking up the effect

Figure 2: Baseline Performance Regressions over 0-20 day Horizons



Notes: this figure plots the estimated β coefficients from our baseline regression 4.2 up to 20-day horizon ($T = 20$), using the value weighted performance variable as the regressand, measured in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Transactions"). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. The shaded area denotes the 90% confidence band, It is based on robust standard errors, using two-way clustering at the month and the client level.

of trading volume (driving both connections and performance), in panel 4b, we extend this analysis and double-sort our sample using the within-variation both in connections and in trading volume. The performance difference in high and low connectivity months is approximately the same irrespective of whether the client's trading volume is high or low, and thereby the majority of positive trading performance continues to coincide with high connectivity months.

Moreover, we assess the persistence of the effect of connections and gradually increase the trading horizon up to 20 days ($T = 20$) while re-estimating our baseline regression 4.2. In Figure 2, we present the 20 estimated βs, using the value weighted performance measure, together with the 90% confidence bands. We find that the effect peaks at the 5-day horizon, but we still find that client connectivity significantly affects performance at the 10-day horizon. The effect then gradually decays, with the point estimate reaching zero at the 18-day horizon.

According to our model, the intuition is as follows. The precision of clients information varies with time. If she systematically chooses her counterparty from a wider set of dealers whenever her private information is more precise (i.e. requests quotes from more dealers), the adverse selection her regular dealer faces is smaller. Therefore, her regular dealer can provide a narrower bid-ask spread in average. This is how the client can recover the (reputational or search) cost of multiple quotes. At the same time, as periods with multiple connections will coincide with periods with more precise private information, this mechanism provides a positive relationship between clients' connections and subsequent trading performance. Note that our theory does not imply causality between connections and performance in any direction. Instead, both higher performance and higher connectedness are caused by more private information.

4.2.2 Decomposing Trading Performance into Transaction and Anticipation Components

Given our baseline results, we now explore the channels through which client connectivity is related to trading performance. Specifically, we test whether more connected clients may perform better because they get better deals compared to other clients trading around the same time (transaction component) and/or because they can better anticipate future price changes over the coming trading days (anticipation component). Our mechanism does not have strong predictions on the earlier, but requires the latter effect to be present. To this end, we estimate two modified versions of our baseline specification (4.2) with the trading performance measure replaced with the anticipation and transaction components (3.2). Table 5 shows the decomposition results for the 5-day value-weighted performance measure.

Our results show that a client, when more connected, tends to perform significantly better in each component. When more connected, she tends to trade at a more favourable price and to the direction of future price movements. Quantitatively, we find that the anticipation component has much larger role in the overall higher performance of clients when they are more connected. In particular, less than 20% of our baseline effect is explained by the transaction component, with the anticipation component (0.4bp) being responsible for the majority of our baseline effect (0.5bp).

So far, our empirical results provide support for Hypotheses 1 and 3. We now turn to Hypothesis 2, namely, that the relationship between connectivity and trading performance is stronger when price volatility is higher.

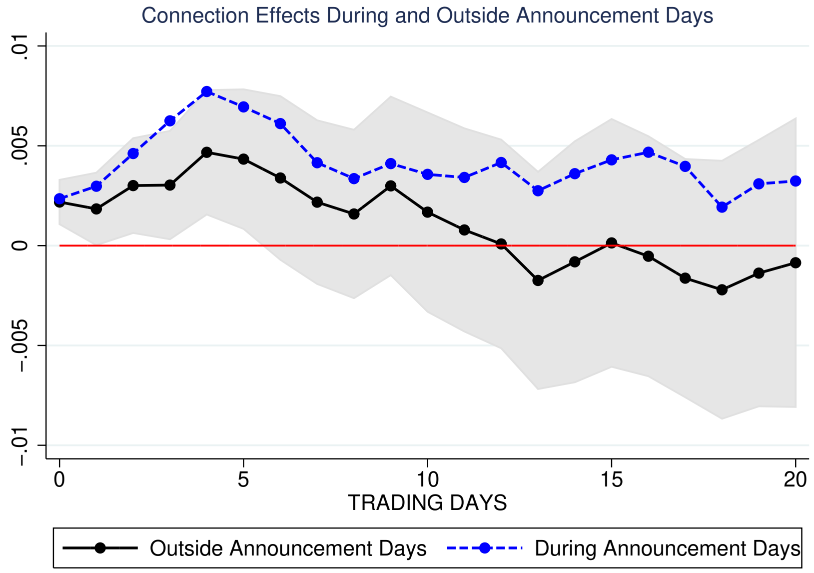
4.2.3 The Role of Macroeconomic Announcements

Since Fleming and Remolona (1999); Brandt and Kavajecz (2004), there has been ample empirical evidence on the effect of scheduled macroeconomic announcements on government bond prices and volatility. Green (2004) finds that the informational role of trading increases following announcements, indicating that the release of public information raises the level of information asymmetry in the government bond market. His evidence suggests that some market participants have an advantage at processing the newly arrived information. Motivated by this empirical evidence and the related theoretical literature (Pasquariello and Vega, 2007; Kondor, 2012), we now explore whether the relation between connections and performance is different during the arrival of public information. Our baseline analysis relies on UK monetary policy announcements and the release of the consumer price index. Policy announcements include the publication of the quarterly inflation report, the policy interest rate decision of the Monetary Policy Committee (MPC) and the release of the minutes (Table 11).[15][16] Out of the 1470 trading days in our sample, we end up with 196 trading days that coincide with news about the policy interest rate and inflation. In the spirit of our analysis above, we compute two sets of monthly performance measures for each of our client: one that is based on all announcement days, and another based on all other trading days without announcements. Accordingly, we extend our baseline regression 4.2. and estimate the following model:

[15]See Gerko and Rey (2017) for further details on the institutional arrangements of the UK and US monetary policy decision making process.

[16]In Subsection 6.4, we show that our results are robust to using alternative definitions of macroeconomic announcements.

Figure 3: Performance Regressions over 0-20 day Horizons: Trading Days With and Without Release of Macroeconomic News



Notes: the black line plots the estimated β coefficients from the regression 4.3 up to 20-day horizon ($T = 20$), using the value weighted performance variable (based on non-announcement days) as the regressand, measured in basis points. The blue line plots the sum of the estimated coefficients $\rho + \beta$, representing the effect on announcement days. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Transactions"). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. The shaded area denotes the 90% confidence band associated with the estimated β coefficients, It is based on robust standard errors, using two-way clustering at the month and the client level.

$$\begin{aligned} Performance^T_{i,t,p} = \rho \times D^{Ann}_{i,p} \times ClientConnections_{i,t} \\ + \beta \times ClientConnections_{i,t} + X_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t,p}, \end{aligned} \tag{4.3}$$

where $D^{Ann}_{i,p}$ is a dummy variable taking value 1 if the performance measure is based on trading days with macroeconomic announcements and 0 otherwise. The term ρ is the coefficient of interest which measures whether connectedness has differential effect on performance during announcements, compared to non-announcement days. Table 12 and 13 show that the effect of client connections on trading performance is substantially stronger on trading days of scheduled inflation or interest rate announcements. For example, the point estimate 0.0032 in Table 12 suggests that the effect of trading with an additional dealer on the 3-day performance is twice as strong on an announcement day than on a trading day without announcements, with the difference being highly statistically significant.

To assess the persistence of the effect, we gradually increase the trading horizon up to 20 days ($T = 20$), and re-estimate our regression model 4.3. The black line in Figure 3 represents the 20 estimated βs associated with non-announcement days, using the value weighted performance measure, together with the 90% confidence bands. The blue line shows the effect of connectedness on announcement days. We find that, at all trading horizons, the relationship between connections and trading performance is stronger during macroeconomic announcements than on non-announcement days. This is consistent with Hypothesis 2.

4.3 Aggregate Connections and the Yield Curve

Having presented evidence on the positive relationship of a client's connections and her individual performance, now we turn to the implications of our mechanism to price discovery. Specifically, we test whether time-variation in aggregate client-dealer connections in the market can explain variation in the level and slope of the yield curve.

To estimate the relationship between day-to-day variation in yields and aggregate connections, we first construct an aggregate measure of connections defined as the total number of unique client-dealer connections on a given trading day. We then examine the relationship between changes in aggregate connections and changes in the absolute value of the 5-year yield and of the term spread, defined as the difference between the 25-year and 1-year yields.[17]

Results are summarised in Table 17. Our specification draws on the extensive literature on the relation between price changes and trading volume (Karpoff, 1987; Bessembinder and Seguin, 1993; Chan and Fong, 2000; Malinova and Park, 2011). Following this literature, we also include volume as a control in our regressions.

We find a statistically strong relationship between daily changes in aggregate connections and absolute deviations in yields levels (Top Panel) as well as deviations in the term spread (Bottom Panel). We find that both connections and volume are significant when including them separately in the regression, as shown by Columns 1–2 in Table 17. However, most of the explanatory power of volume disappears once we include both variables in the regression, while connections continue to be strongly significant (Column 3).[18] From a theoretical point of view, this is not surprising as volume more likely captures large liquidity trades as well as informational trades.

[17]We focus on changes in the absolute value of yields, because informationally intensive trading days, captured by aggregate connections, can coincide with both positive and negative price changes.

[18]Table 29 in the Appendix shows similar results for the 10-year and 25-year yields as well.

To show that our result is not mechanically driven by simply more clients being present in the market on high information days, we include as control the total number of unique clients (Tauchen and Pitts, 1983). As shown by Column 5, we find that it is, effectively, the changes in the total number of dealer connections per client (and not the changing number of clients) that drive day-to-day yield changes. Overall, these results suggest the aggregate connections are an important conduit by which new (private and public) information impounds in prices, consistent with Hypothesis 4.

5 Applications

Having established that client connections serve well as a proxy for private information in dealer markets, one can use this proxy to empirically investigate a number of long-standing issues in the finance literature.

In this section, we turn our attention to two questions in particular. First, in Subsection 5.1, we explore the leakage of information from dealers to their preferred clients. In particular, we present suggestive evidence that dealers learn from their informed clients and pass this information to their subsidiaries.[19]

Second, in Subsection 5.2, we are interested in the the nature of private information in treasury markets. This is an intriguing topic as this market is often viewed as a market with little role for private information. We show evidence that information proxied by connections is partially related to future order-flow, i.e. more connected clients better predict the order-flow intermediated by the dealers they trade with.

5.1 Information Leakages

In this part, with the help of our specially detailed data set, we investigate the information leakage from dealers to their preferred clients.

We capture the special relationship between dealers and some of their clients as follows. For each dealer, we are able to distinguish between trading accounts that perform a market-making function (trading primarily with clients, executing large number of transactions, participating in primary auctions) from trading accounts that correspond to other, client-like arms of the given dealer bank (trading primarily with other dealers, executing lower number of transactions, e.g. asset-manager arms). We refer to these

[19] In a related set of papers, Maggio, Franzoni, Kermani, and Sommavilla (2019); Barbon, Maggio, Franzoni, and Landier (2017) shows that brokers in corporate bond and stock markets pass on order-flow information to the clients that they have had a strong trading relationships.

accounts as the given dealer's subsidiaries. We then test whether dealers' subsidiaries perform better when the given dealer trades with a larger proportion of high-connection clients.

Separating dealers from their subsidiaries also provides a useful source of variation to disentangle *inventory* effects from *information* effects. That is, focusing on the trades of dealers (instead of the trades of their subsidiaries) would make it more difficult to tell whether dealers trade (on the inter-dealer broker market) into the same direction as their informed clients simply because dealers want to offset these trades thereby reducing inventory risk or because dealers indeed learn profitable information from their informed clients. Focusing on dealers' subsidiaries, instead of dealers, helps us better isolate informational effects from inventory management – an integral part of market making (see Goldstein and Hotchkiss (2019) and references herein).

To use our connectivity measure to proxy the informativeness of client order flow that a given dealer faces, we construct the following measure for each dealer i on day t:

$$InfShare_{i,t} = \frac{Vol^H_{i,t}}{Vol^L_{i,t} + Vol^H_{i,t}}, \tag{5.1}$$

where $Vol^H_{i,t}$ and $Vol^L_{i,t}$ are the trading volume of dealer i with clients whose connectivity on day t is higher and lower, respectively, than their sample average. Again, we rely purely on the time-series variation in connectivity when sorting clients, so that half the daily observations of each client contribute to measure $Vol^H_{i,t}$ and the other half of observations contribute to measure $Vol^L_{i,t}$.

We use measure 3.1 to compute the daily trading performance of dealers' subsidiaries, $SubsidPerformance_{i,t}$. To test whether subsidiaries perform better when their dealers trade with more connected clients, we estimate the following daily panel regression for each dealer i and day t:

$$SubsidPerformance^T_{i,t} = \beta \times InfShare_{i,t} + X_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}, \tag{5.2}$$

where α_i and μ_t are firm and time fixed effects; and $X_{i,t}$ includes four control variables such as the number of transactions and trading volume of dealers' subsidiaries as well as the dealers. The main coefficient of interest in 5.2 is β which captures the relation between dealers' enhanced interaction with high-connection clients and the performance of dealers' subsidiaries.

The results in Table 14 show that when a dealer's clientele goes from low connectivity to high-connectivity, then the trading performance of the dealer's subsidiaries improves by 1-2 basis points over the 1-2 day horizon. Interestingly, there is no effect at the 0-day horizon suggesting that we are not simply picking up that dealers' subsidiaries get better information about bid-ask spreads; the information they learn is informative about imminent changes in the yield curve over the coming days. It is worth noting that these results are not just detecting that larger/better dealers who tend to attract more informed (and connected) clients might have higher performing subsidiaries: the inclusion of dealer fixed effects absorbs this type of time-invariant heterogeneity across dealers.

Moreover, it is worth emphasising that we control for dealers' trading volume as well, so the estimated β is not simply picking up that dealers' increased trading volume itself (Campbell, Grossman, and Wang 1993; Wang 1994) conveys important information about yields to dealers' subsidiaries while potentially being correlated with the quality of the dealer's clientele. The composition of dealers' clientele is indeed the factor that determines how dealers' subsidiaries perform (against other dealers), suggesting that subsidiaries obtain the information that their dealers learn from informed clients.

Two possible concerns with this conclusion are that (i) the information, which increases the profitability of dealers' subsidiaries, might originate from other dealers that these clients are connected to, or (ii) it might originate from the subsidiaries themselves. The information could then be passed on to these clients' dealers who disseminate it across their other clients – and this is what regression 5.2 might be picking up.

To reinforce our story, that the information flows from dealers to their subsidiaries and not the other way around, we perturb our research design by adding two control variables to regression 5.2. First, we build on measure 5.1 to compute the average informativeness of all other dealers that the given subsidiary is connected to (excluding the given subsidiary's own dealer from this average measure). The constructed variable ("InfShare of OtherDealers") is aimed to control for the first identification concern. Second, we include as a control the number of dealer connections of subsidiaries ("Subsid Connections") to address the second concern. Table 15 shows that these additional controls are statistically insignificant and their inclusion in the regression makes little difference to the coefficient on $InfShare_{i,t}$. We interpret these results that the information is in fact flowing from dealers to their subsidiaries and not the other way around.

In addition, we relate this analysis to that of Maggio, Franzoni, Kermani, and Sommavilla (2019) which focuses on the cross-sectional heterogeneity in the eigenvalue-centrality of stock-market brokers to study information diffusion. The main premise of their paper is that more central brokers gather and disseminate more information than less central brokers do. This begs the question of whether the eigenvalue-centrality of a dealer in our application proxies the time-variation in the share of connected clients in the dealer's total client base, measured by $InfShare_{i,t}$. To check this, we include the eigenvalue-centrality of dealers as a control in regression 5.2. Table 16 shows that dealer centrality is statistically insignificant and including it in the regression makes little difference to the coefficient on $InfShare_{i,t}$.

5.2 The Nature of Private Information

In this part, we investigate the nature of private information clients' connections may proxy. Our starting point is the empirical literature (Fleming and Remolona 1999; Evans and Lyons 2002; Brandt and Kavajecz 2004; Menkveld, Sarkar, and van der Wel 2012) on the role of order flow in driving prices in various dealer markets. This literature observed that agents who have information about the future order flow in these markets can use this information profitably.

Motivated by this, we proceed in three steps. First, as presented in Subsection 5.2.1, we propose a measure of co-movement of the composition of client's orders with the future aggregate order flow of a given group of clients. The idea is that whenever this measure is positive, the client, intentionally or by chance, is effectively front-running that group of clients. We test whether this measure identifies profitable trades. We indeed find that whenever the duration composition of a client's trade is similar to that of all the other clients in subsequent days, her performance is higher. Second, as presented in Subsection 5.2.2, we connect our baseline results to order flow information: we show that whenever a client is more connected, the composition of her trades tend to be more similar to the group of clients in subsequent days who are served by the same dealer. We also show that a client who is a regular counterparty of the given dealer can predict the composition of the order flow better. This suggests that dealers have an important role in disseminating order flow information towards their own, regular clients. Finally, as presented in Subsection 5.2.3, we also show that all our findings are stronger for the group of clients who drive our baseline performance result.

5.2.1 Measuring Co-movement between Client Trades and Future Order Flow

Our proposed measure aims to capture whether a client trades in the same direction as other clients in the subsequent trading days. First, we partition all transactions in K

equal-sized segments based on the modified duration of all traded gilts. We then compute the net trading position of client i, on day d, in duration segment k, $W_{i,d,k}$.[20] We then calculate the cumulative net trading position of group g between days $d+1$ and $d+T$ in duration segment k, denoted by $W^g_{d+T,k}$. The identity of group g will play an important role in section 5.2.2 where we identify the group whose order flow connected clients can forecast. For now, we set g for the group of all the clients in the market. Our daily covariance measure, $\Psi^{T,g}_{i,d}$, is then computed as follows:

$$\Psi^{T,g}_{i,d} = \frac{1}{K}\sum_{k=1}^{K}\left(W_{i,d,k} - \frac{1}{K}\sum_{k=1}^{K}W_{i,d,k}\right)\left(W^g_{d+T,k} - \frac{1}{K}\sum_{k=1}^{K}W^g_{d+T,k}\right). \tag{5.3}$$

When $\Psi^{T,g}_{i,d}$ is high, the given client tends to concentrate her orders in the same segment as group g in the subsequent T days. Figure 8 in the Appendix provides a pictorial illustration of measure 5.3.[21] Given this measure, we first check whether it is profitable to guess right the segments of the yield curve where future demand pressure will be concentrated. For each client i we partition the trading days into two sets, $p \in \{Low, High\}$. A day is in the high (low) set, if $\Psi^{T,g}_{i,d}$ for the given day is larger (smaller) than the median in the full sample. Then we estimate the following regression:

$$Performance^T_{i,t,p} = \gamma \times Q_{i,p} + \delta_{i,t} + \varepsilon_{i,t,p}, \tag{5.4}$$

where $Performance^T_{i,t,p}$ is the version of our baseline performance measure (3.1) which is aggregated only over set p of trading days in months t for client i. $Q_{i,p}$ is a dummy taking value 1 if the performance measure of client i is based on high-covariance trading days and 0 if it is based on low-covariance days. The term $\delta_{i,t}$ is a client-month fixed effect. Table 7 summarises the results. Panel A shows the results when the covariance measure uses the cumulative order flow of the market ($g = Total$) at the 3-day horizon (columns 1-3) or 5-day horizon (columns 4-6). For both cases, we compute the turnover-weighted performance measures at the 1-, 3- and 5-day horizons. We find that the trading performance of a client can be 2-3bp higher on high covariance days, i.e. predicting the order flow of the market is profitable. Panel B shows the results when the covariance measure uses the cumulative order flow of the subset of the market that is intermediated

[20] Clients' net positions corresponds to their order flow in this market, as client-dealer trades are initiated by clients.

[21] The scatter plot in this illustration (bottom panel of Figure 8) can in fact be estimated using our dataset. Figure 4 below will do exactly that.

by the dealers that the given client is connected to ($g = Own$). We find that if a client can predict this subset of the aggregate order flow, it is still profitable with performance being 1-2bp higher on high covariance days.

5.2.2 Connected Clients Predict the Order Flow

Let us return to our baseline result that the time-variation in a client's number of connections is a proxy for her level of private information. In this section, we provide evidence that this private information is on the duration composition of the future order flow of certain group of other clients, as measured by our covariance measure 5.3. In this case, we expect that the covariance measure of a given client in a given month tends to be higher when this client is more connected. Hence, we compute monthly averages of $\Psi^{T,g}_{i,d}$, denoted as $\Psi^{T,g}_{i,t}$ to estimate the following panel regression:

$$\Psi^{T,g}_{i,t} = \phi \times ClientConnections_{i,t} + X_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}, \tag{5.5}$$

where the terms on the right-hand-side are identical to our baseline specification 4.2.

Panel A of Table 8 shows the results for the total order flow ($g = Total$). While we see a positive relationship between the time-variation in connections and the duration decomposition of client's current trades at that of the aggregate order flow, this relationship is not strong. Instead, in Panel B and C, we decompose aggregate order flow as follows. We isolate the part of the aggregate order flow that is intermediated through the dealers which a given client is connected to ($g = Own$) from the part that goes through all the other dealers that the given client is not connected to ($g = Non - Own$). We further decompose the Own measure based on whether the given client has a more regular relationship with a dealer ($g = Regular$), distinguishing it from other client-dealer connections that are relatively new ($g = New$). We regard a client-dealer connection regular if the client traded with the given dealer in the current as well as in the previous month; whereas we regard a connection new if a client traded with the dealer in the current month but not in the previous month.[22]

Perhaps the most intriguing finding of this part is in Panel B of Table 8. It decomposes the aggregate private client flow into the part that is intermediated by those

[22]Note that, by the additivity of covariance, our measure is additive in the following sense:

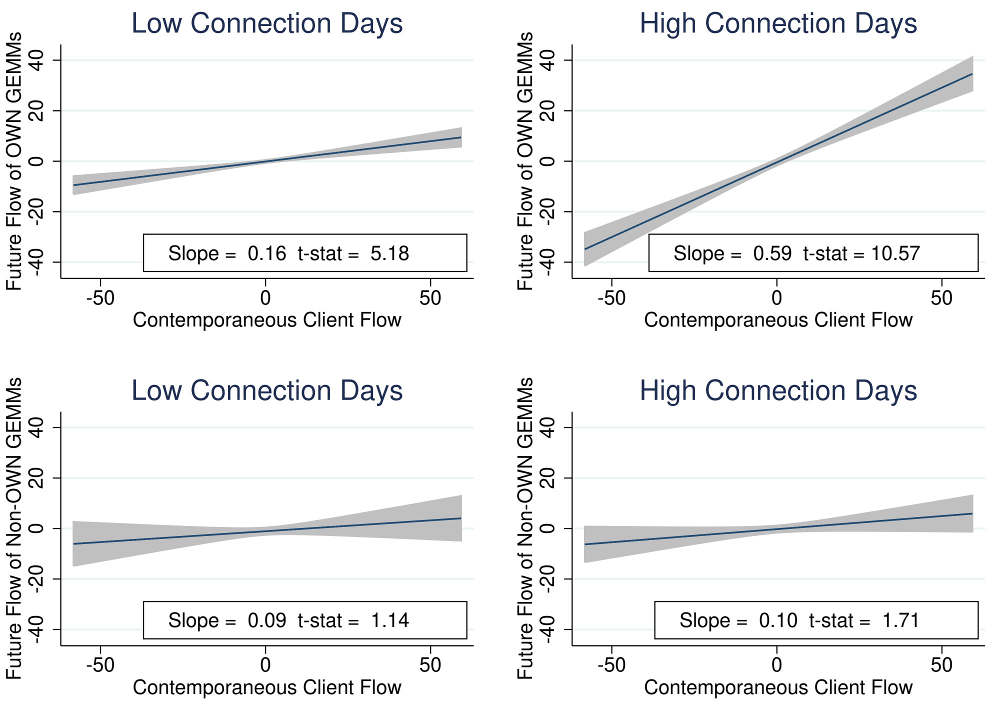
$$YC^{T,Total}_{i,t} = YC^{T,Own}_{i,t} + YC^{T,Non-Own}_{i,t} = YC^{T,Regular}_{i,t} + YC^{T,New}_{i,t} + YC^{T,Non-Own}_{i,t}. \tag{5.6}$$

This property helps the interpretation of our results.

dealers that the given client is connected to (Columns 1-3) in contrast with the order flow that is channelled through dealers that the client is not connected to (Columns 4-6). We find that it is the covariance with *Own* dealer order flow that correlates with the client's connectivity, and the effects for $Non - Own$ dealer order flow are economically and statistically insignificant. Our interpretation is that dealers, intentionally or unintentionally, disseminate information about future orders towards (some of) their clients. We have little evidence on the exact mechanism. In principle, dealers' private information on their clients expected orders in the subsequent days might be revealed accidentally by the dealers' quotes. Or it might be that there is an intentional information flow from dealers to their best clients helping dealers to keep these clients (Maggio, Kermani, and Song, 2017). Indeed, Panel C shows that higher client connectivity predicts more the order flow intermediated by *Regular* dealers and predict less the order flow that goes through newly connected dealers.

We provide further evidence on the importance of dealers in forecasting future flows, by using our dataset at a more disaggregated level. First, for each client i, we compute the daily net trading position in each duration bucket k (5.3). Second, we compute the future net trading position (cumulated over the subsequent T days) aggregated across all clients that trade with those dealers who are connected with client i in the given month. We refer to this as the future flow of Own dealers. We also compute the aggregate future client flow that is intermediated by those dealers who are not connected with client i in the given month. We refer to this as the future flow of Non-own dealers. Consistent with the evidence in Table 8, we expect to find that the relationship between client order flow and future flow of Own dealers should be higher in those months when the client has higher level of connectivity. In turn, we do not expect the relationship between client order flow and future flow Non-own dealers to be different in high and low connectivity months. Figure 4 shows the estimated regressions slopes from four separate linear regressions. The top (bottom) row shows the relationship between client flow and future flow of Own (Non-own) dealers. The right (left) column shows the relationship between the flows when the client is more (less) connected compared to its own average. We find that client flows co-move more strongly with Own dealer flows than with Non-own dealer flows, reflected by the statistically insignificant slope coefficients in the bottom row of Figure 4. Importantly, we find the strongest co-movement between client flows and future Own dealer flows when clients are more connected. This is shown by the regression coefficient in the top right panel (0.59) being almost four times larger than the slope in the top left panel (0.16).

Figure 4: Contemporaneous Client Order Flow and Future Aggregate Order Flow: the Roles of Connectivity and Client-Dealer Relations



Notes: Each panel shows the estimated regression slope (with associated 90% confidence interval) that corresponds to the relationship between contemporaneous order flow of an individual client i and the aggregate cumulative future order flow of all clients. The units of observation are daily and duration-specific net trading positions of a given client i that are regressed against daily and duration-specific aggregate cumulative future net trading positions of all clients that are intermediated by all the dealers that client i is connected to (top row) and by all other dealers that client i is not connected to (bottom row). The left (right) column looks at the flow relationships in those months when client i is in the bottom (top) quartile of client connectedness (using within variation). The axes are measured in £000,000s.

Overall, the results of this and the previous subsections suggest that the nature of private information, proxied by client connectivity, is related to the order flow intermediated by the given client's dealers.

5.2.3 Variation in Clients' Performance Sensitivity

The results in this Section suggest that when a client is more connected her trades predict better the future order flows of his client. But is this behind the higher performance in this period? To establish a connection to our baseline results, we now exploit the heterogeneity across clients masked by Figure 2.

For some of our clients there is a stronger co-movement between connections and performance while for others this comovement could be much weaker. Our interpretation is that not all clients are in the market to profit from short-term bets based on private information. In this section, we explore variation in the sensitivity of client's trading performance to connections to provide additional evidence to our narrative. In particular, we enforce the insights that (1) time-variation in connections is a proxy for private information on the duration composition of future order flows, (2) this information is disseminated by dealers to their own clients only, (3) regular clients tend to be more the recipients of this information.

We proceed in two steps. First, we re-estimate our baseline regression (3.1) for each client separately, and then sort the clients based on the their estimated β coefficients. We define a dummy variable, $D_i^{H\beta}$, which takes value 1 if the client's estimated β is above the median ('high-β clients') and takes value 0 otherwise ('low-β clients'). Second, we extend the empirical model 5.5, by adding the interaction term $D_i^{H\beta} \times ClientConnections_{i,t}$ to it, and estimate the following panel-regression:

$$\begin{aligned}\Psi_{i,t}^{T,g} &= \rho \times D_i^{H\beta} \times ClientConnections_{i,t} \\ &\quad + \beta \times ClientConnections_{i,t} + X_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t,p},\end{aligned} \tag{5.7}$$

where ρ is the coefficient of interest which measures whether the effect of connections on the covariance measure is higher amongst those clients whose trading performance is more sensitive to connections ($D_i^{H\beta}$). Tables 9–10 confirm that the effect between connections and the ability to predict the order flow is significantly stronger amongst high-β clients compared to low-β clients. This is true for the aggregate order flow just as well, as for our decompositions. This group can forecast better the order flow of their own clients and the effect is stronger for clients who are regular than for those who are new comers to the particular dealer.

6 Sensitivity Analysis and Further Explorations

6.1 Performance Regressions Using Daily Data

Our baseline results were based on data at monthly frequency. This was to reduce measurement noise and also to avoid oversampling those clients who trade actively, possibly on most trading days. However, one concern might be that monthly averaging introduces problems of time aggregation which makes it difficult to accurately measure the dynamics of client-dealer connections. To address this, we re-estimate most of our regressions on daily data. Interestingly, the time-series variation in client connections continues to be substantial when we go to daily frequency. Table 22 shows that the standard deviation of connections is about 1.5 when using only the within-variation, i.e. the average client frequently changes the number of dealers that she trades across trading days.

Results for our performance regressions, presented in Table 23, are similar to our monthly regression results – though seem less persistent with the connectivity effect peaking at the four-day horizon. In Figure 7, we plot the estimated connectedness coefficients for longer horizons along with the interaction effects of monetary policy announcements, corroborating the increased strength of the performance-connectedness relationship during the arrival of public information.

We also test at daily frequency whether information about the duration composition of future market order flow increases performance.[23] Table 24 shows that when the duration composition of the client's order flow has high covariance with that of the market one day (Table 25a) or three days (Table 25b) after the client traded, then the client's short-run performance increases by 2-3bp. We argued that clients are less likely to predict the total market order flow than the order flow intermediated by the client's own dealers. We use our daily data to test whether such partial but more accessible information about the order flow still significantly profitable, by changing out dummy variable to $G = Own$. Table 25 shows that the estimated coefficients for $Q^G_{i,p}$ are roughly halved compared to Table 24, suggesting that predicting the order flow intermediated by the client's own dealers is, on average, about half as valuable for making profits than predicting the order flow of the whole market.

We now connect our results about client performance (Table 23) and about order flow predictions (24–25), and estimate whether clients can better predict their own dealer's order flow on trading days with unusually high connectedness. Unlike in our monthly specification, note that we define the client's own dealers with reference to the given trading day: own dealers are the ones that the client traded with on the day of the trade as well as during the past 10 trading days. The more accurate measurement of the timing of client-dealer connections allows us to better separate the time-variation in connectedness from the formation of client-dealer relationships, thereby assessing whether client connectedness is part of the information acquisition process or it is merely an instrument of concealing information.

[23] To show this, we estimate the daily variant of regression 5.4, $Performance^T_{i,t} = \gamma \times Q^G_{i,t} + X_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t}$, where $Q^G_{i,p}$ (with $G = Total$) is a dummy taking value 1 if the performance measure of client i is based on high-covariance trading days and 0 if it is based on low-covariance days (Formula 5.3), with respect to the total market order flow.

To assess this, we turn to our covariance-connectedness regressions (5.5). Specifically, we estimate whether client connectedness predicts the covariance of the client's order flow with the whole market. As shown by Panel a of Table 26, we find only weak evidence for this, consistent with our monthly results. Moreover, as presented in Panel b of Table 26, we estimate whether higher connections predict higher covariance with the order flow of those dealers that the client traded with on the given trading day *as well as* during the preceding 10 trading days (columns 1-3). This is contrasted with the results when the covariance between the client's order flow and its dealer's is based on those dealers that the client only traded with on the given trading day, but not during the preceding 10 trading days (columns 4-6). The fact that more connected clients cannot predict the order flow of newly connected dealers (only that of the regular dealers) suggests that connectedness is not instrumental in the information acquisition process, but it merely helps transform information into better performance.

6.2 Predicting Changing Yield Curve vs Noise

We also explore whether a client's T-day performance on a trade can be high because the client can better predict changes in the shape of the yield curve, or because the client can better predict changes in the distance of individual gilt yields (pricing error) from an otherwise unchanged yield curve. We compute the decomposition for each transaction j as follows:

$$\ln\left(P^T\right)-\ln\left(P_j^\star\right) \approx \underbrace{\left[\ln\left(M^T\right)-\ln\left(M^0\right)\right]}_{Curve-Shift}+\underbrace{\left\{\left[\ln\left(P^T\right)-\ln\left(M^T\right)\right]+\left[\ln\left(M^0\right)-\ln\left(P_j^\star\right)\right]\right\}}_{Pricing-Error}, \tag{6.1}$$

where M^0 and M^T are the end of 0-day and T-day prices implied by standard yield curve models (Nelson and Siegel, 1987; Svensson, 1994; Hu, Pan, and Wang, 2013).

Figure 6 shows the decomposition of our baseline connection effect into the component of yield curve forecasts and noise forecasts as specified by (6.1). Looking at the two figures together, there is some evidence that the yield curve component is stronger than the pricing-error component. Note, that if pricing-errors are not persistent beyond a day, than the transitional component in (3.2) is expected to be close to the pricing-error component in (6.1).

6.3 An Alternative Performance Measure

Our performance measure compares the transaction price with the future market price of the security. Whether a client liquidates her position at that future price, or holds on to it, does not influence our measure. That is, our performance measure might not correspond to realised profits. This is in contrast to the bulk of previous empirical work on over-the-counter markets which measures performance as the return on dealers' round-trip transactions. The reason why we do not follow that approach is that our focus is not on the performance of dealers – who trade very frequently and tend to finish their day with small net positions – but on clients. In our sample, clients trade for heterogeneous reasons. In aggregate, they tend to persistently accumulate positions as they ultimately purchase most of the issued securities by the DMO. Individually, some trade frequently, some buy and hold. Some might aim for profit by turning over their portfolio quickly, others might aim instead to acquire their desired positions at a favourable price. Our performance measure is neutral to the objective of the client. If the client manages to buy at a low price or sell at a high price compared to the price in the subsequent period, we measure that as a high value transaction.

Still, as a robustness test, we construct a second performance measure which measures realised profit in a given month directly, building on the average-cost-approach of inventory valuation. In particular, for each client i, gilt a, month m, we compute:

$$R_{i,a,m} = \left[\ln \left(\frac{\sum_{j^S=1}^{J^S_{i,a,m}} P_{i,a,j^S} Q_{i,a,j^S}}{\sum_{j^S=1}^{J^S_{i,a,m}} Q_{i,a,j^S}} \right) - \ln \left(\frac{\sum_{j^B=1}^{J^B_{i,a,m}} P_{i,a,j^B} Q_{i,a,j^B}}{\sum_{j^B=1}^{J^B_{i,a,m}} Q_{i,a,j^B}} \right) \right] \times \min \left[\sum_{j^S=1}^{J^S_{i,a,m}} Q_{i,a,j^S}, \sum_{j^B=1}^{J^B_{i,a,m}} Q_{i,a,j^B} \right], \tag{6.2}$$

where $J^p_{i,a,m}$, with $p = S, B$, denote the total number of monthly, gilt-specific sale and buy transactions, respectively, while P_{i,a,j^p} and Q_{i,a,j^p} corresponds to the price and quantity of transaction j^p. We then compute the weighted average of $R_{i,a,m}$ across gilts (using the client's monthly trading volume in gilt a as weights) to obtain a realized profit measure at the client-month level. We then re-estimate our baseline specification with our within-month realized profit measure on the right hand side.

As the first panel of Table 6 shows, while more connections of a given client are associated with higher realised profit, this relationship is not significant in the full sample. However, if we focus on those client-month observations when the given client trades frequently (i.e. more than the median number of transactions), the relationship is significant. Our interpretation is that our within-month realised profit measure captures high value trades only for those clients and for those periods when the client trades a lot within a month.

6.4 Expanding the Definition of Macroeconomic Announcements

We found that the connected-performance relationships is stronger during days that coincide with the announcement of macroeconomic news (Section 4.2.3). We now test whether this result is robust to expanding the definition of macroeconomic announcements. To do so, we include additional announcement days in our analysis such as the release of UK real activity indicators (unemployment, average earnings, manufacturing production and GDP) as these indicators have been found to strongly affect the government bond markets in the US (Fleming and Remolona 1999). Moreover, we also add the days of the release of US FOMC statements and minutes, as recent evidence showed strong effects of US monetary policy shocks affecting global financial markets (Miranda-Agrippino and Rey (2015); Gerko and Rey (2017)). This leaves us with 422 trading days that coincide with macroeconomic announcements. Tables 27–28 show that the results are similar to our baseline.

6.5 Aggregate Connections and Asymmetric Information

As an addition test, we check whether variation in aggregate connections may explain aggregate measures of asymmetric information as well. Following Jankowitsch, Nashikkar, and Subrahmanyam (2011); Friewald, Jankowitsch, and Subrahmanyam (2012), we compute for each day the following transaction-based measure that has been shown to be particularly suitable for OTC markets:

$$\Omega = \sqrt{\frac{1}{\sum_{k=1}^{K} v_k} \cdot \sum_{k=1}^{K} \left(\log\left(p_k\right) - \log\left(p^\star\right)\right)^2 \cdot v_k}, \tag{6.3}$$

where K_t is the number of observed transactions for all bonds on a given trading day, p_k (for $k = 1, \ldots, K$) is the transaction price, $p^\star$ is the closing mid-price of the corresponding bond, and v_k is the size of each trade. Measure 6.3 is the root mean squared difference between the traded prices and the respective market-wide valuation. Results are summarised in Table 18 showing a similar picture to Table 17. A 1% increase in total connections is associated with a 1-1.5% increase in aggregate price dispersion, depending on the exact regression specification. Again, the effect of connections seems to dominate the effects of trading volume and number of clients in the market.

6.6 The Centrality of Dealers

As mentioned above, our first application on information leakages is related to the recent work Maggio, Franzoni, Kermani, and Sommavilla (2019) which focuses on the heterogeneity in the eigenvalue-centrality of stock-market brokers to study information diffusion. In contrast, our study is distinct as we identify informed transactions by focusing on the time-variation in the connections of clients (who wish to hide private information). Nevertheless, we revisit their analysis in our dataset and estimate their monthly regression model (1) for client i, dealer j and month t:

$$Performance^T_{i,j,t} = \beta \times DealerCentrality_{j,t} + X_{j,t} + \alpha_{i,t} + \mu_j + \varepsilon_{i,j,t}, \qquad (6.4)$$

where $Performance^T_{i,j,t}$ is the performance of client i against dealer j in month t; $X_{j,t}$ includes dealer-specific controls such as trading volume and average trade size as in Maggio, Franzoni, Kermani, and Sommavilla (2019). Following their level of disaggregation, the data is now collapsed at the client-dealer-month level, so that we can include a client-time fixed effect, $\alpha_{i,t}$ (compared to the client-month level as in our baseline specification 4.2).

Table 30 in Appendix B.5 shows the results. Qualitatively, we can replicate their results (Panel A of their Table 3): dealer centrality and volume are positively related to client performance, while average trade size is negatively related. However, the statistical significance of dealer centrality substantially weakens once volume is included in the regression. This might be because the effect they identified for stock market brokers is harder to identify for the UK government bond market where dealers are more homogeneous, leading to more limited cross-sectional variation in their centrality.

Finally, we explore whether our measure of client connections captures information that is independent of that captured by dealer centrality. The unconditional correlation between the two variables at the client-month-dealer level is -0.03, suggesting that they are close to being orthogonal to each other. To further check this, we exclude client-time fixed effects, and include both client connections and dealer centrality in regression 6.4.[24]

In Table 31 shows the results for the value-weighted performance measure from 0- to 5-day horizon. In Panel A, we only include dealer centrality in the regression. In Panel B, we only include client connections. We find that both variables are statistically

[24] Recall that client connections vary only at the client-time level, hence client-time fixed effects need excluding.

significant, depending on the performance horizon. More importantly, as shown in Panel C, we find that the estimated coefficients are similar, when including both variables in the same regression, to those obtained when including them in separate regressions (Panel A–B). This reinforces the point that the informational content of client connections is distinct from that of dealer centrality.

7 Summary and Conclusion

Our paper provided evidence from the UK government bond market that clients better predict future price movements when they have more dealer connections compared to periods when they have fewer connections. This effect is stronger around macroeconomic announcements. We also showed that innovations in the slope and level of the yield curve are associated with days of higher aggregate connections in the market. Based on these findings, we argued that time-variation in client connections serves as an empirical proxy for time-variation in private information. We also presented two applications using this proxy. We found evidence suggesting that dealers leak the information deduced from their client base to their subsidiaries. We also established that part of the private information identified by connections is related to the maturity structure of the order-flow the given client's dealer is receiving in subsequent days.

These results have several implications. First, our results highlight the relevance of financial network formation to the price discovery process in government bond markets. While the literature has extensively studied the role of private information and aggregate order flow in determining yield curve dynamics, we find that a better understanding of the network structure can sharpen our understanding of the price discovery process in these markets. Second, while a number of recent papers have studied the core-periphery structure of OTC markets (primarily focusing on the cross-sectional characteristics of dealer-client relationships), our results emphasize the dynamic and endogenous nature of networks. Third, slow trade execution is often regarded as optimal because it minimizes price impact, thereby helping to hide private information (Kyle, 1985). We find that trade execution with multiple primary dealers could serve a similar purpose, suggesting that splitting trades over time and across dealers may be substitutable.

A clear caveat of our approach is that to calculate clients' connections, a detailed, transaction level data-set, including the identities of market participants, is required. As the trend seems to be that such data-sets are becoming increasingly accessible to the academic community, we expect that our approach opens up new avenues to better understand the role of private information in financial markets.

References

Afonso, G., A. Kovner, and A. Schoar (2014): "Trading partners in the interbank lending market," Staff Reports 620, Federal Reserve Bank of New York. 3.3.2

Babus, A., and P. Kondor (2018): "Trading and Information Diffusion in Over-the-Counter Markets," *Econometrica*, 86(5), 1727–1769. 3

Back, K., P. Collin-Dufresne, V. Fos, T. Li, and A. Ljungqvist (2018): "Activism, Strategic Trading, and Liquidity," *Econometrica*, 86(4), 1431–1463. 9

Balduzzi, P., E. J. Elton, and T. C. Green (2001): "Economic News and Bond Prices: Evidence from the U.S. Treasury Market," *The Journal of Financial and Quantitative Analysis*, 36(4), 523–543. 1

Barbon, A., M. D. Maggio, F. Franzoni, and A. Landier (2017): "Brokers and Order Flow Leakage: Evidence from Fire Sales," NBER Working Papers 24089, National Bureau of Economic Research, Inc. 19

Benos, E., and F. Zikes (2018): "Funding constraints and liquidity in two-tiered OTC markets," *Journal of Financial Markets*. 7

Bessembinder, H., and P. J. Seguin (1993): "Price Volatility, Trading Volume, and Market Depth: Evidence from Futures Markets," *The Journal of Financial and Quantitative Analysis*, 28(1), 21–39. 4.3

Bonacich, P., and P. Lloyd (2001): "Eigenvector-like measures of centrality for asymmetric relations," *Social Networks*, 23(3), 191 – 201. 10

Boulatov, A., T. Hendershott, and D. Livdan (2013): "Informed Trading and Portfolio Returns," *Review of Economic Studies*, 80(1), 35–72. 1

Brancaccio, G., D. Li, and N. Schurhoff (2017): "Learning by Trading: The Case of the U.S. Market for Municipal Bonds," mimeo. 3

Brandt, M. W., and K. A. Kavajecz (2004): "Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve," *Journal of Finance*, 59(6), 2623–2654. 1, 4.2.3, 5.2

Campbell, J. Y., S. J. Grossman, and J. Wang (1993): "Trading Volume and Serial Correlation in Stock Returns," *The Quarterly Journal of Economics*, 108(4), 905–939. 5.1

Chakravarty, S., H. Gulen, and S. Mayhew (2004): "Informed Trading in Stock and Option Markets," *Journal of Finance*, 59(3), 1235–1258. 1

Chan, K., and W.-M. Fong (2000): "Trade size, order imbalance, and the volatility-volume relation," *Journal of Financial Economics*, 57(2), 247–273. 4.3

Chen, H.-L., N. Jegadeesh, and R. Wermers (2000): "The Value of Active Mutual Fund Management: An Examination of the Stockholdings and Trades of Fund Managers," *Journal of Financial and Quantitative Analysis*, 35(03), 343–368. 14

Cohen, L., C. Malloy, and L. Pomorski (2012): "Decoding Inside Information," *The Journal of Finance*, 67(3), 1009–1043. 1

Collin-Dufresne, P., and V. Fos (2015): "Do Prices Reveal the Presence of Informed Trading?," *Journal of Finance*, 70(4), 1555–1582. 1

DMO (2011): "A guide to the roles of the DMO and Primary Dealers in the UK government bond market," Discussion paper, Debt Management Office. 2.1

Duarte, J., and L. Young (2009): "Why is PIN priced?," *Journal of Financial Economics*, 91(2), 119–138. 1

Duffie, D., S. Malamud, and G. Manso (2009): "Information Percolation With Equilibrium Search Dynamics," *Econometrica*, 77(5), 1513–1574. 3

Easley, D., N. M. Kiefer, M. O'Hara, and J. B. Paperman (1996): "Liquidity, Information, and Infrequently Traded Stocks," *The Journal of Finance*, 51(4), 1405–1436. 1

Easley, D., and M. O'Hara (1987): "Price, trade size, and information in securities markets," *Journal of Financial Economics*, 19(1), 69 – 90. 4.2.1

Elton, Edwin J, e. a. (1993): "Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios," *Review of Financial Studies*, 6(1), 1–22. 14

Evans, M., and R. Lyons (2002): "Order Flow and Exchange Rate Dynamics," *Journal of Political Economy*, 110(1), 170–180. 5.2

Fleming, M. J., and E. M. Remolona (1999): "Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information," *Journal of Finance*, 54(5), 1901–1915. 1, 4.2.3, 5.2, 6.4

Friewald, N., R. Jankowitsch, and M. G. Subrahmanyam (2012): "Illiquidity or credit deterioration: A study of liquidity in the US corporate bond market during financial crises," *Journal of Financial Economics*, 105(1), 18–36. 6.5

Gabrieli, S., and C.-P. Georg (2014): "A network view on interbank market freezes," Discussion paper. 1

Garleanu, N., and L. H. Pedersen (2013): "Dynamic Trading with Predictable Returns and Transaction Costs," *Journal of Finance*, 68(6), 2309–2340. 9

Gerko, E., and H. Rey (2017): "Monetary Policy in the Capitals of Capital," *Journal of the European Economic Association*, 15(4), 721–745. 15, 6.4

Glosten, L. R., and P. R. Milgrom (1985): "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders," *Journal of Financial Economics*, 14(1), 71–100. 1, 2.2, A

Goldstein, M. A., and E. S. Hotchkiss (2019): "Providing liquidity in an illiquid market: Dealer behavior in US corporate bonds," *Journal of Financial Economics*, Forthcoming. 5.1

Golosov, M., G. Lorenzoni, and A. Tsyvinski (2014): "Decentralized Trading with Private Information," *Econometrica*, 82(3), 1055–1091. 3

Green, T. C. (2004): "Economic News and the Impact of Trading on Bond Prices," *Journal of Finance*, 59(3), 1201–1234. 1, 4.2.3

Hagstromer, B., and A. Menkveld (2019): "Information Revelation in Decentralized Markets," *The Journal of Finance*, (Forthcoming). 4

Hendershott, T., D. Li, D. Livdan, and S. Norman (2017): "Relationship Trading in OTC Markets," CEPR Discussion Papers 12472, C.E.P.R. Discussion Papers. 3.3.2

Hendershott, T., D. Livdan, and N. Schurhoff (2015): "Are institutions informed about news?," *Journal of Financial Economics*, 117(2), 249–287. 1

Hollifield, B., A. Neklyudov, and C. Spatt (2017): "Bid-Ask Spreads, Trading Networks, and the Pricing of Securitizations," *Review of Financial Studies*, 30(9), 3048–3085. 1, 3.3.2

Hortacsu, A., and J. Kastl (2012): "Valuing Dealers' Informational Advantage: A Study of Canadian Treasury Auctions," *Econometrica*, 80(6), 2511–2542. 1

Hu, G. X., J. Pan, and J. Wang (2013): "Noise as Information for Illiquidity," *The Journal of Finance*, 68(6), 2341–2382. 6.2

Jankowitsch, R., A. Nashikkar, and M. G. Subrahmanyam (2011): "Price dispersion in OTC markets: A new measure of liquidity," *Journal of Banking and Finance*, 35(2), 343–357. 6.5

Johnson, T. L., and E. C. So (2018): "A Simple Multimarket Measure of Information Asymmetry," *Management Science*, 64(3), 1055–1080. 1

Kacperczyk, M., and E. Pagnotta (2019): "Chasing Private Information," *Review of Financial Studies*, Forthcoming. 2

Kacperczyk, M., C. Sialm, and L. Zheng (2005): "On the Industry Concentration of Actively Managed Equity Mutual Funds," *Journal of Finance*, 60(4), 1983–2011. 14

Karpoff, J. (1987): "The Relation between Price Changes and Trading Volume: A Survey," *Journal of Financial and Quantitative Analysis*, 22(01), 109–126. 4.3

Kondor, P. (2012): “The More We Know about the Fundamental, the Less We Agree on the Price,” *Review of Economic Studies*, 79(3), 1175–1207. 4.2.3

Kyle, A. (1985): “Continuous Auctions and Insider Trading,” *Econometrica*, 53(6), 1315–35. 1, 2.2, 9, 7

Li, D., and N. Schurhoff (2019): “Dealer Networks,” *The Journal of Finance*, 74(1), 91–144. 1

Maggio, M. D., F. Franzoni, A. Kermani, and C. Sommavilla (2019): “The relevance of broker networks for information diffusion in the stock market,” *Journal of Financial Economics*. 1, 3.3.1, 4.2.1, 19, 5.1, 6.6, 6.6, 16, **??**

Maggio, M. D., A. Kermani, and Z. Song (2017): “The value of trading relations in turbulent times,” *Journal of Financial Economics*, 124(2), 266 – 284. 10, 3.3.2, 5.2.2

Malinova, K., and A. Park (2011): “Trading Volume in Dealer Markets,” *Journal of Financial and Quantitative Analysis*, 45(06), 1447–1484. 4.3

Mascio, R. D., A. Lines, and N. Y. Naik (2017): “Alpha Decay,” mimeo. 9

Menkveld, A., A. Sarkar, and M. van der Wel (2012): “Customer Order Flow, Intermediaries, and Discovery of the Equilibrium Risk-Free Rate,” *Journal of Financial and Quantitative Analysis*, 47(04), 821–849. 5.2

Merrick, J., N. Y. Naik, and P. K. Yadav (2005): “Strategic trading behavior and price distortion in a manipulated market: anatomy of a squeeze,” *Journal of Financial Economics*, 77(1), 171–218. 4.2.1

Meulbroek, L. K. (1992): “An Empirical Analysis of Illegal Insider Trading,” *Journal of Finance*, 47(5), 1661–99. 2

Miranda-Agrippino, S., and H. Rey (2015): “US Monetary Policy and the Global Financial Cycle,” NBER Working Papers 21722, National Bureau of Economic Research, Inc. 6.4

Nelson, C. R., and A. F. Siegel (1987): “Parsimonious Modeling of Yield Curves,” *The Journal of Business*, 60(4), 473–489. 6.2

Pasquariello, P., and C. Vega (2007): “Informed and Strategic Order Flow in the Bond Markets,” *Review of Financial Studies*, 20(6), 1975–2019. 1, 4.2.3

Roll, R., E. Schwartz, and A. Subrahmanyam (2010): “O/S: The relative trading activity in options and stock,” *Journal of Financial Economics*, 96(1), 1–17. 1

Svensson, L. E. (1994): “Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994,” NBER Working Papers 4871, National Bureau of Economic Research, Inc. 6.2

Tauchen, G., and M. Pitts (1983): “The Price Variability-Volume Relationship on Speculative Markets,” *Econometrica*, 51(2), 485–505. 4.3

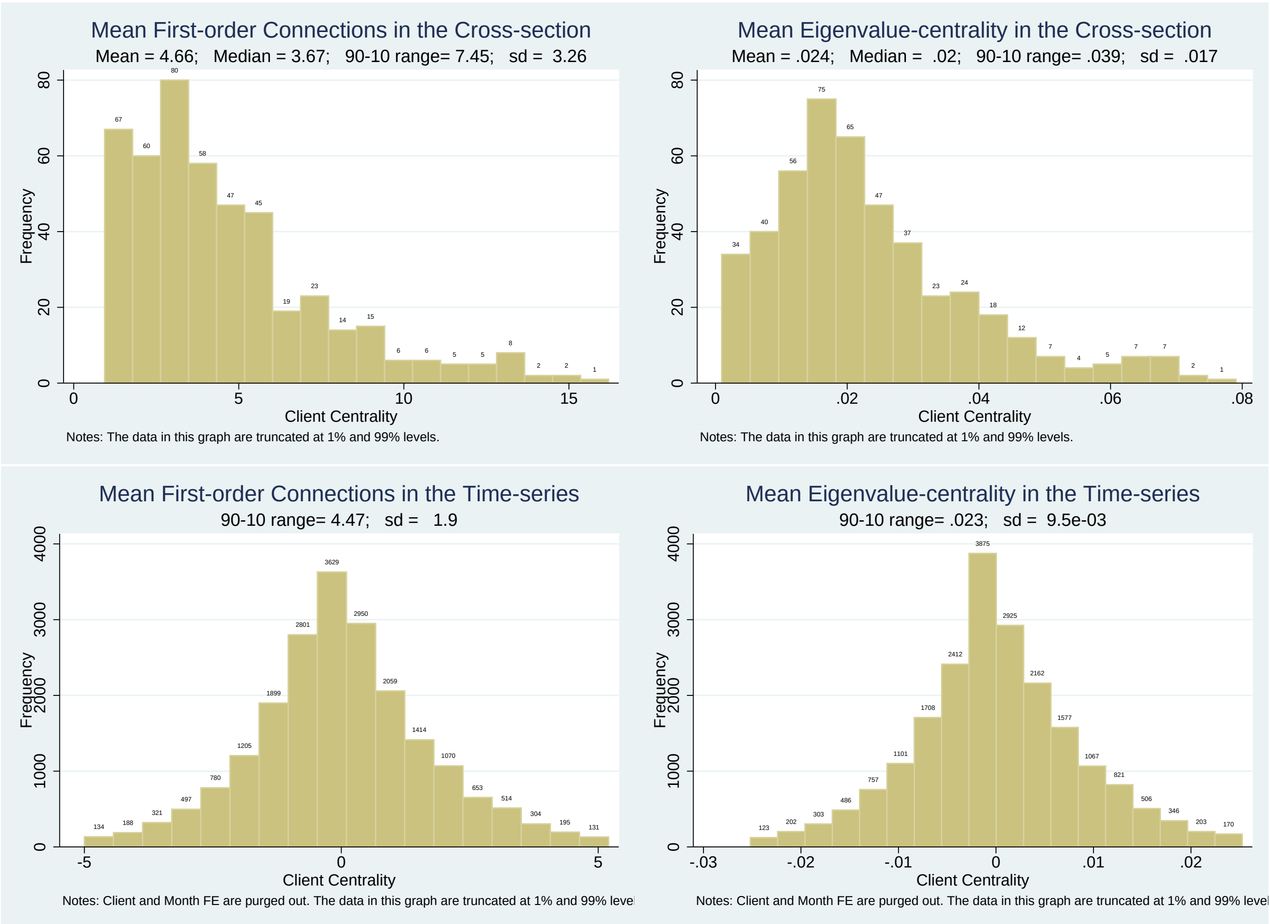
Valseth, S. (2013): "Price discovery in government bond markets," *Journal of Financial Markets*, 16(1), 127–151. 1

Wang, J. (1994): "A Model of Competitive Stock Trading Volume," *Journal of Political Economy*, 102(1), 127–168. 5.1

8 Figures and Tables

8.1 Summary Statistics

Figure 5: Time-series and Cross-sectional Variation in Connectivity



Notes: these figures summarize the time-series and cross-sectional variation in our first-order (left column) and eigenvalue-centrality (right column) measures. The top row plots the distribution of mean client connectedness. To construct the bottom row, we first run a panel regression to purge out client and month fixed effects ($Connections_{i,t} = \alpha_i + \mu_t + \varepsilon_{i,t}$), and plot the distribution of the residuals ($\varepsilon_{i,t}$).

Table 2: Summary Statistics – Month-Client Level

(a) All Clients

	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	Median	p10	p90	sd	N
First Order Connection	5.434	4	1	12	4.086	21,170
Eigenvalue-centrality	0.0275	0.0230	0.00483	0.0588	0.0206	21,170
Transaction Number	86.70	19	4	184	274.8	21,170
$log(Volume)$	17.29	17.50	13.40	20.82	2.821	21,170
Average Monthly Duration	8.621	8.288	3.379	14.03	4.232	21,170

(b) Well Connected vs Less Connected Clients [Across-Client Variation]

	Below Median Connections		Above Median Connections		Diff.
	Mean	Median	Mean	Median	t-stat
5-day Weighted Performance	-0.0591	0	1.084	0.794	-4.40***
5-day Unweighted Performance	-0.396	-0.0253	1.206	0.888	-3.05***
3-day Weighted Performance	-0.591	-0.235	0.696	0.476	-3.25***
3-day Unweighted Performance	-1.020	-0.502	0.661	0.418	-2.09**

(c) Well Connected vs Less Connected Months [Within-Client Variation]

	Below Median Connections		Above Median Connections		Diff.
	Mean	Median	Mean	Median	t-stat
5-day Weighted Performance	-0.0909	0.297	1.445	0.682	-2.77***
5-day Unweighted Performance	-0.224	0.363	1.326	0.624	-2.98***
3-day Weighted Performance	-0.473	-3.69e-07	0.828	0.424	-3.02***
3-day Unweighted Performance	-0.648	-0.0521	0.449	0.221	-2.72***

Notes: This table reports summary statistics for our baseline sample, covering 2011m10-2017m6, that is collapsed at the month-client level. Panel A reports summary statistics for all clients. Panel B and C report summary statistics on unweighted and volume-weighted performance measures at the 3-day and 5-day horizons, measured in basis points. Panel B differentiates between more connected and less connected clients by placing clients, in each month, into two groups based on whether their first-order centrality measure is below or above the median in the given month. Panel C places each client observation into two groups based on the within-variation of connections, i.e. depending on whether the client's first-order centrality measure is below or above the client's own median centrality measure based on the whole sample. The last column in Panel B and C reports the t-statistics associated with the test of whether performance is different for low and high connectivity clients (Panel B) and for low and high connectivity months (Panel C). Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

8.2 Baseline Results

Table 3: Client Connections and Trading Performance: Baseline Results

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.0019*	0.0033***	0.0039***	0.0048***	0.0051***
Connections	(1.82)	(2.66)	(2.78)	(2.74)	(2.86)
Volume	-0.0004	0.0004	0.0003	-0.0013	-0.0018
	(-0.24)	(0.17)	(0.11)	(-0.43)	(-0.51)
Tran.	0.0031	0.0005	-0.0023	-0.0049	-0.0041
	(1.11)	(0.12)	(-0.53)	(-0.95)	(-0.64)
N	20839	20839	20839	20839	20839
R^2	0.037	0.039	0.036	0.034	0.036
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(a) Turnover-weighted Trading Performance

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.0019*	0.0032***	0.0033**	0.0046***	0.0051***
Connections	(1.97)	(2.76)	(2.25)	(2.75)	(2.97)
Volume	-0.0011	0.0014	0.0028	0.0019	0.0007
	(-0.64)	(0.63)	(1.06)	(0.70)	(0.21)
Tran.	0.0012	-0.0037	-0.0059	-0.0074	-0.0055
	(0.37)	(-0.94)	(-1.34)	(-1.54)	(-1.02)
N	20839	20839	20839	20839	20839
R^2	0.051	0.047	0.044	0.039	0.039
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Unweighted Trading Performance

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on client connections (4.2). The transaction-level data is collapsed at the client-month level. The performance measures are in %-points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 4: Illustrating the Economic Significance of Connections

(a) Single-sorting by Connections

	Average Monthly Volume (in £000s)		Average Monthly 5-day Performance		Decomposition of Gross Performance	
	Mean	Median	Mean	Median	Mean	Median
Low Connection Months	308,000	15,000	-0.274	0.167	-15%	6%
High Connection Months	507,000	53,100	1.279	0.721	115%	94%
					100%	100%

(b) Double-sorting by Connections and Volume

		Average Monthly Volume (in £000s)		Average Monthly 5-day Performance		Decomposition of Gross Performance	
		Mean	Median	Mean	Median	Mean	Median
Low Volume Months	Low Connection Months	156,000	5,230	-0.097	0.00234	-1%	1%
Low Volume Months	High Connection Months	285,000	23,300	1.414	0.00821	40%	20%
High Volume Months	Low Connection Months	466,000	38,100	-0.456	0.00124	-21%	5%
High Volume Months	High Connection Months	742,000	111,000	1.137	0.00645	83%	74%
						100%	100%

Note: This table illustrates the economic significance of the performance-connectedness relationship. Panel *a* splits the sample (at the client-month level) into two groups using single-sorting, based on the within-variation of connections, i.e. the first (second) group contains the observations for those months when the given client had fewer (more) monthly connections compared to its sample average. Panel *b* splits the sample (at the client-month level) into four groups using double-sorting, based on the within-variation of both connections and trading volume. The numbers in bold decompose gross performance (defined as the product of volume and performance) into the contribution of each group. The 5-day performance measures are in basis points.

Table 5: Client Connections and the 5-day Performance: Transaction vs Anticipation Effect

	(1) Baseline	(2) Transaction [Id]	(3) Anticip. [Id]	(4) Transaction [D]	(5) Anticip. [D]
Client	0.0051***	0.0005	0.0044**	0.0011**	0.0037**
Connections	(2.86)	(1.54)	(2.56)	(2.28)	(2.16)
Volume	-0.0018	-0.0006	-0.0012	-0.0014	-0.0004
	(-0.51)	(-1.00)	(-0.32)	(-1.57)	(-0.10)
Tran.	-0.0041	0.0005	-0.0042	-0.0007	-0.0028
	(-0.64)	(0.56)	(-0.66)	(-0.50)	(-0.45)
N	20839	20839	20839	20839	20839
R^2	0.036	0.082	0.034	0.073	0.034
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(a) Turnover-weighted Trading Performance

	(1) Baseline	(2) Transaction [Id]	(3) Anticip. [Id]	(4) Transaction [D]	(5) Anticip. [D]
Client	0.0051***	0.0008**	0.0043**	0.0012***	0.0037**
Connections	(2.97)	(2.39)	(2.54)	(2.74)	(2.28)
Volume	0.0007	-0.0006	0.0012	-0.0012	0.0017
	(0.21)	(-1.11)	(0.38)	(-1.56)	(0.54)
Tran.	-0.0055	0.0004	-0.0058	-0.0003	-0.0052
	(-1.02)	(0.39)	(-1.08)	(-0.19)	(-0.98)
N	20839	20839	20839	20839	20839
R^2	0.039	0.143	0.035	0.116	0.034
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Unweighted Trading Performance

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at the 5-day horizon horizons on client connections (4.2). The decomposition is based on 3.2. The results in Columns (2)-(3) are based on the average transaction price $\overline{P}$ that uses the trades (for the given gilt) in a 3-hour window within the trading day. The results in Columns (4)-(5) are based on $\overline{P}$ being the average transaction price of all trades in the given trading day. The transaction-level data is collapsed at the client-month level. We include as a control the natural logarithm of the pound trade volume of each client and the natural logarithm of the number of monthly transactions. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

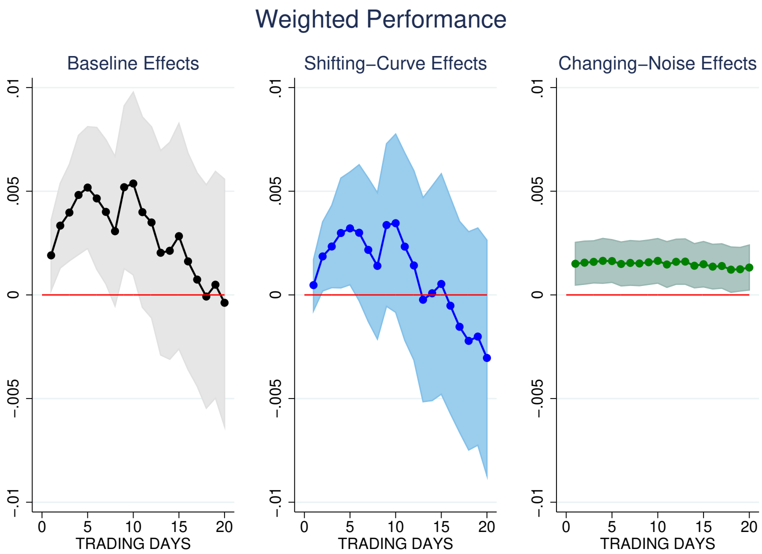
Table 6: Client Connections and Realised Performance

	(1) Full Sample	(2) High Transaction Months
Client	0.0037	0.0058**
Connections	(1.36)	(2.03)
Volume	-0.0017	-0.0034
	(-0.17)	(-0.30)
Tran.	-0.0032	-0.0077
	(-0.63)	(-1.02)
N	15242	9721
R^2	0.059	0.073
Time FE	Yes	Yes
Client FE	Yes	Yes

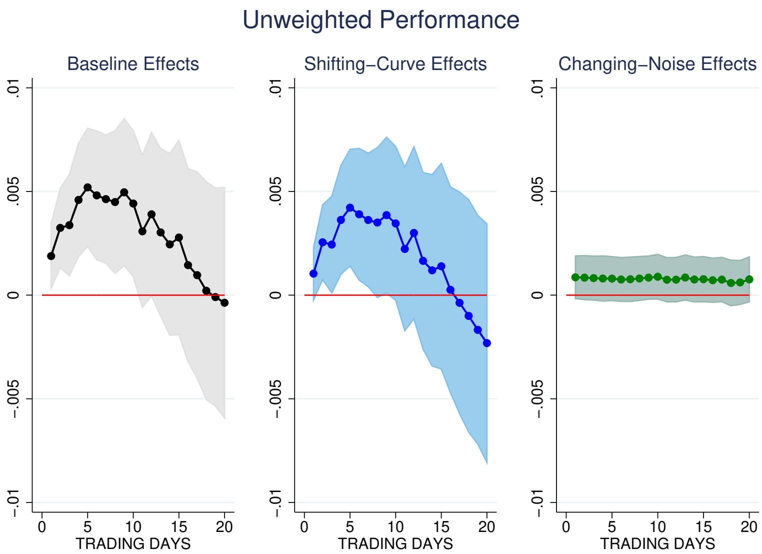
Notes: this table regresses the realised trading performance (6.2) on connections. The transaction-level data is collapsed at the client-month level. We include as a control the natural logarithm of the pound trade volume of each client and the natural logarithm of the number of monthly transactions. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Figure 6: Decomposing the Baseline Performance into Yield Curve Shifting and Changing Noise Effects

(a) Decomposing Turnover Weighted Performance



(b) Decomposing Unweighted Performance



Notes: this figure plots the estimated β coefficients from variants of our baseline regression 4.2, where we use three measures of performance according to 6.1: (i) the baseline performance measure, (ii) the yield-curve shift component, and (iii) the noise component. We estimate the regression up to 20-day horizon ($T = 20$), using the turnover weighted (unweighted) performance variable as the regressand in panel A (panel B). The performance measure is in %. We include as a control the natural logarithm of the pound trade volume of each client and the natural logarithm of the number of monthly transactions. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. The shaded area denotes the 90% confidence band, It is based on robust standard errors, using two-way clustering at the month and the client level.

8.3 Connectivity and Predicting the Order Flow

Table 7: Trading Performance on High and Low Covariance Days

	3-day Covariance			5-day Covariance		
	1-day Perf	3-day Perf	5-day Perf	1-day Perf	3-day Perf	5-day Perf
γ	0.0368***	0.0207**	0.0001	0.0354***	0.0299**	0.0156
	(4.70)	(2.26)	(0.01)	(3.05)	(2.38)	(1.14)
N	35130	34944	34770	35130	34944	34770
R^2	0.520	0.503	0.501	0.523	0.512	0.502

(a) Covariance with the Total Order Flow

	3-day Covariance			5-day Covariance		
	1-day Perf	3-day Perf	5-day Perf	1-day Perf	3-day Perf	5-day Perf
γ	0.0189***	0.0136**	0.0034	0.0162**	0.0169**	0.0143*
	(3.95)	(2.33)	(0.51)	(2.41)	(2.11)	(1.69)
N	34346	34300	34198	34346	34300	34198
R^2	0.520	0.508	0.495	0.529	0.509	0.500

(b) Covariance with the Order Flow of Own GEMMs

Notes: this table regresses the turnover-weighted trading performance at different time horizons on a dummy taking value 1 if the performance measure is based on high-covariance trades and 0 if it is based on low-covariance days (5.4). The two panels differ in terms of how the covariance is computed: Panel (a) computes the covariance with the total order flow ($g = Total$); Panel (b) computes the covariance with the aggregate order flow intermediated by the client's own GEMMs ($g = Own$). The transaction-level data is collapsed at the client-month level. The performance measures are in percentage points. We winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 8: Client Connectivity and Covariance with the Aggregate Order Flow

	1-Day	3-Day	5-Day
Client	0.0062*	0.0021	-0.0007
Connections	(1.68)	(0.48)	(-0.16)
Volume	-0.0083	-0.0076	-0.0031
	(-1.04)	(-0.93)	(-0.36)
Tran.	0.0132	0.0283**	0.0226
	(0.99)	(2.03)	(1.59)
N	20289	20284	20279
R^2	0.033	0.031	0.028
Time/Client FE	Yes/Yes	Yes/Yes	Yes/Yes

(a) Total Client Order Flow

	Own GEMMs' Order Flow			Rest of GEMMs' Order Flow		
	1-Day	3-Day	5-Day	1-Day	3-Day	5-Day
Client	0.0109**	0.0109**	0.0111**	0.0027	0.0003	-0.0040
Connections	(2.22)	(2.21)	(2.23)	(0.56)	(0.07)	(-0.89)
Volume	-0.0048	-0.0126	-0.0083	-0.0085	-0.0143*	-0.0053
	(-0.48)	(-1.22)	(-0.85)	(-1.00)	(-1.79)	(-0.62)
Tran.	0.0226	0.0201	0.0086	0.0071	0.0295**	0.0192
	(1.50)	(1.17)	(0.56)	(0.47)	(2.26)	(1.53)
N	18932	18929	18926	18932	18929	18926
R^2	0.036	0.032	0.033	0.031	0.034	0.031
Time/Client FE	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes

(b) Total Client Order Flow via Own GEMMS vs Non-own GEMMs

	Regular Client-GEMM Connections			New Client-GEMM Connections		
	1-Day	3-Day	5-Day	1-Day	3-Day	5-Day
Client	0.0078*	0.0125**	0.0085*	0.0069*	-0.0003	0.0071
Connections	(1.71)	(2.60)	(1.81)	(1.73)	(-0.08)	(1.65)
Volume	0.0041	0.0020	0.0036	0.0016	-0.0093	-0.0139
	(0.52)	(0.25)	(0.41)	(0.16)	(-0.97)	(-1.43)
Tran.	0.0114	0.0007	-0.0054	0.0033	0.0265*	0.0223
	(0.81)	(0.05)	(-0.38)	(0.23)	(1.77)	(1.51)
N	18932	18929	18926	18932	18929	18926
R^2	0.037	0.033	0.033	0.028	0.028	0.030
Time/Client FE	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes

(c) Total Client Order Flow via Own GEMMS: Regular vs New Connections

Notes: this table regresses different versions of the covariance measure 5.3 on our connectivity measure and controls (5.5). The transaction-level data is collapsed at the client-month level. The performance measures are in basis points. In Panel (a) the outcome variable is the client's covariance with the aggregate client order flow in the market. In Panel (b) the outcome variable is the client's covariance with the aggregate order flow intermediated by the dealers that the client is connected to (columns 1-3), and by all other dealers (columns 4-6). In Panel (c), the outcome variable is the client's covariance with the aggregate order flow intermediated by the dealers that the client is regularly trades with (columns 1-3) and by all other dealers that the client trades with in the given month but not in the previous month. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 9: Client Connectivity and Covariance with the Aggregate Order Flow: The Role of High Performance Sensitivity Clients

	Total Market Order Flow		
	1-Day	3-Day	5-Day
Connections	-0.0055	-0.0036	-0.0048
	(-0.96)	(-0.56)	(-0.83)
Connections$\times D_{\beta}^{H}$	0.0166***	0.0071	0.0030
	(4.04)	(1.54)	(0.57)
Volume	-0.0086	-0.0077	-0.0032
	(-1.08)	(-0.94)	(-0.37)
Tran.	0.0134	0.0284**	0.0227
	(1.02)	(2.05)	(1.59)
N	20289	20284	20279
R^2	0.034	0.031	0.028
Time FE	Yes	Yes	Yes
Client FE	Yes	Yes	Yes

Notes: this table regresses the covariance measure 5.3 (with the aggregate market order flow) on our connectivity measure interacted with a dummy indicating high-β clients as well as other controls (5.7). The transaction-level data is collapsed at the client-month level. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 10: Client Connectivity and Covariance with the Aggregate Order Flow: The Role of High Performance Sensitivity Clients

	Own GEMMs' Order Flow			Rest of GEMMs' Order Flow		
	1-Day	3-Day	5-Day	1-Day	3-Day	5-Day
Connections	-0.0026	0.0004	0.0016	-0.0025	-0.0042	-0.0081
	(-0.37)	(0.06)	(0.25)	(-0.42)	(-0.63)	(-1.34)
Connections$\times D^H_\beta$	0.0230***	0.0204***	0.0197***	0.0073	0.0044	-0.0003
	(4.68)	(3.52)	(3.22)	(1.36)	(0.94)	(-0.07)
Volume	-0.0050	-0.0128	-0.0085	-0.0086	-0.0144*	-0.0054
	(-0.51)	(-1.24)	(-0.86)	(-1.01)	(-1.81)	(-0.63)
Tran.	0.0230	0.0203	0.0089	0.0072	0.0296**	0.0193
	(1.53)	(1.20)	(0.58)	(0.48)	(2.28)	(1.54)
N	18932	18929	18926	18932	18929	18926
R^2	0.037	0.032	0.034	0.031	0.034	0.031
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes	Yes

(a) Total Client Order Flow via Own GEMMS vs Non-own GEMMs

	Regular Client-GEMM Connections			New Client-GEMM Connections		
	1-Day	3-Day	5-Day	1-Day	3-Day	5-Day
Connections	-0.0037	0.0041	-0.0005	0.0005	-0.0081	0.0027
	(-0.50)	(0.61)	(-0.07)	(0.10)	(-1.62)	(0.52)
Connections$\times D^H_\beta$	0.0182***	0.0201***	0.0165***	0.0127**	0.0067	0.0110*
	(4.09)	(3.53)	(3.09)	(2.38)	(1.12)	(1.69)
Volume	0.0039	0.0019	0.0034	0.0015	-0.0095	-0.0140
	(0.49)	(0.23)	(0.39)	(0.15)	(-0.98)	(-1.43)
Tran.	0.0117	0.0010	-0.0052	0.0034	0.0267*	0.0224
	(0.83)	(0.07)	(-0.36)	(0.24)	(1.79)	(1.53)
N	18932	18929	18926	18932	18929	18926
R^2	0.038	0.033	0.033	0.028	0.028	0.030
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes	Yes

(b) Total Client Order Flow via Own GEMMS: Regular vs New Connections

Notes: this table the covariance measure 5.3 on our connectivity measure interacted with a dummy (D^H_β) indicating high-β clients as well as other controls (5.7). The transaction-level data is collapsed at the client-month level. In Panel (a) the outcome variable is the client's covariance with the aggregate order flow intermediated by the dealers that the client is connected to (columns 1-3), and by all other dealers (columns 4-6). In Panel (b), the outcome variable is the client's covariance with the aggregate order flow intermediated by the dealers that the client is regularly trades with (columns 1-3) and by all other dealers that the client trades with in the given month but not in the previous month. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

8.4 Macroeconomic Announcements

Table 11: Economic Announcements

Announcement	Source	Cumulated Number of Announcement Days
Panel A: Core Announcements		
UK Inflation Report	Bank of England	
UK MPC Minutes	Bank of England	127
UK MPC Decision	Bank of England	
UK Inflation Rate	ONS	196
Panel B: Additional Announcements		
UK Earnings/Unemployment	ONS	
UK Manufacturing	ONS	356
UK GDP	ONS	
US FOMC Minutes	Federal Reserve	422
US FOMC Statement	Federal Reserve	

Notes: The table lists the major macroeconomic announcements that our analysis focuses on. Panel A lists the announcements related to UK nominal variables that we use for our benchmark analysis. Panel B includes additional announcements related to UK real variables and US monetary policy decisions. The third column denotes the cumulated number of trading days in our sample that coincide with macroeconomic announcements. In total, our sample includes 1470 trading days.

Table 12: Turnover-weighted Performance: Trading Days With and Without Inflation & Monetary Policy News

	1-day	2-day	3-day	4-day	5-day
Connections	0.0018*	0.0030**	0.0030*	0.0047**	0.0043**
	(1.67)	(2.09)	(1.84)	(2.47)	(2.04)
Connections$\times ANN'$	0.0011	0.0016*	0.0032***	0.0030**	0.0026*
	(1.66)	(1.98)	(3.30)	(2.49)	(1.71)
Volume	0.0013	0.0023	0.0038	0.0026	0.0017
	(0.71)	(1.01)	(1.36)	(0.90)	(0.48)
Tran.	0.0003	-0.0023	-0.0082*	-0.0108**	-0.0067
	(0.10)	(-0.51)	(-1.68)	(-2.05)	(-1.12)
N	34483	34483	34483	34483	34483
R^2	0.025	0.026	0.024	0.023	0.023
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

Notes: this table regresses the value-weighted trading performance at different time horizons on client connections and connections interacted with a dummy that takes value 1 when the trading days coincide macroeconomic announcements. The transaction-level data is collapsed at the client-month-dummy level, i.e. for each month and each client we compute two sets of performance measures and controls, one set based on announcements days the other set based on trading days without announcements. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 13: Turnover-weighted Performance: Trading Days With and Without Inflation & Monetary Policy News

	1-day	2-day	3-day	4-day	5-day
$\beta_{ANN'}$	0.0011	0.0015*	0.0025**	0.0024*	0.0022
	(1.62)	(1.77)	(2.29)	(1.91)	(1.44)
N	27174	27174	27174	27174	27174
R^2	0.503	0.509	0.516	0.520	0.521
Client-Time FE	Yes	Yes	Yes	Yes	Yes

Notes: this table regresses the value-weighted trading performance at different time horizons on client connections and connections interacted with a dummy that takes value 1 when the trading days coincide macroeconomic announcements. The transaction-level data is collapsed at the client-month-dummy level, i.e. for each month and each client we compute two sets of performance measures, one set based on announcements days the other set based on trading days without announcements. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

8.5 Dealers' Subsidiaries

Table 14: Dealers' Informed Clientele and the Performance of Dealers' Subsidiaries: Baseline

	0-day	1-day	2-day	3-day
InfShare	-0.0012	0.0155**	0.0231**	0.0191
	(-0.28)	(2.69)	(2.35)	(1.43)
DealerVolume	-0.0015	-0.0051*	-0.0026	0.0007
	(-1.18)	(-2.03)	(-0.71)	(0.15)
DealerTran.	0.0017	0.0084	0.0005	-0.0024
	(0.72)	(1.54)	(0.06)	(-0.28)
SubsidVolume	-0.0019*	-0.0033*	-0.0013	-0.0019
	(-2.04)	(-1.99)	(-0.58)	(-0.75)
SubsidTran.	-0.0005	0.0045*	0.0020	0.0001
	(-0.33)	(1.96)	(0.72)	(0.01)
N	20950	20950	20950	20950
R^2	0.072	0.075	0.075	0.074
Time FE	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes

Notes: this table shows the results for regression 5.2, which regresses the value-weighted trading performance of dealers' subsidiaries at different time horizons on our informativeness measure (5.1). The transaction-level data is collapsed at the firm-day level. The performance measures are in %-points. We include as controls the natural logarithm of the pound trade volume of dealers ("DealerVolume") and subsidiaries ("SsubsidVolume") the natural logarithm of the number of daily transactions of dealers ("DealerTran.") and subsidiaries ("SubsidTran."). To reduce noise, we winsorise the sample at the 1%-level. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 15: Dealers' Informed Clientele and the Performance of Dealers' Subsidiaries: Adding More Controls

	0-day	1-day	2-day	3-day
InfShare	-0.0016	0.0153**	0.0219**	0.0171
	(-0.36)	(2.72)	(2.12)	(1.24)
Subsid Connections	-0.0000	-0.0006	0.0001	0.0021
	(-0.05)	(-0.42)	(0.07)	(0.94)
InfShare of OtherDealers	-0.0064	-0.0028	-0.0224	-0.0390
	(-0.75)	(-0.17)	(-0.77)	(-1.16)
DealerVolume	-0.0015	-0.0051*	-0.0026	0.0008
	(-1.18)	(-2.04)	(-0.70)	(0.17)
DealerTran.	0.0017	0.0084	0.0005	-0.0022
	(0.72)	(1.54)	(0.07)	(-0.26)
SubsidVolume	-0.0020*	-0.0034*	-0.0013	-0.0019
	(-2.06)	(-2.01)	(-0.59)	(-0.75)
SubsidTran.	-0.0005	0.0051*	0.0019	-0.0019
	(-0.29)	(2.09)	(0.67)	(-0.43)
N	20950	20950	20950	20950
R^2	0.072	0.075	0.075	0.074
Time FE	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes

Notes: this table shows the results for regression 5.2, which regresses the value-weighted trading performance of dealers' subsidiaries at different time horizons on our informativeness measure (5.1). The transaction-level data is collapsed at the firm-day level. The performance measures are in %-points. We include as controls the connections of the subsidiary ("Subsid Connections"), the informativeness of client order flow of other dealers that the subsidiary is connected to ("InfShare of OtherDealers"), the natural logarithm of the pound trade volume of dealers ("DealerVolume") and subsidiaries ("SubsidVolume") the natural logarithm of the number of daily transactions of dealers ("DealerTran.") and subsidiaries ("SubsidTran."). To reduce noise, we winsorise the sample at the 1%-level. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 16: Dealers' Informed Clientele and the Performance of Dealers' Subsidiaries: Controlling for Dealers' Eigenvalue Centrality

	0-day	1-day	2-day	3-day
InfShare	-0.0011	0.0157**	0.0234**	0.0192
	(-0.26)	(2.71)	(2.37)	(1.44)
Dealer	-0.0489	-0.0765	-0.1194	-0.0524
Eig. Centrality	(-1.37)	(-1.15)	(-1.51)	(-0.46)
DealerVolume	-0.0013	-0.0049*	-0.0023	0.0008
	(-1.07)	(-1.95)	(-0.64)	(0.18)
DealerTran.	0.0031	0.0105*	0.0038	-0.0009
	(1.53)	(2.05)	(0.46)	(-0.10)
SubsidVolume	-0.0019*	-0.0033*	-0.0011	-0.0018
	(-2.01)	(-1.97)	(-0.54)	(-0.73)
SubsidTran.	-0.0000	0.0053**	0.0033	0.0006
	(-0.00)	(2.36)	(0.99)	(0.16)
N	20950	20950	20950	20950
R^2	0.0719	0.0753	0.0754	0.0743
Time FE	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes

Notes: this table shows the results for regression 5.2, which regresses the value-weighted trading performance of dealers' subsidiaries at different time horizons on our informativeness measure (5.1). The transaction-level data is collapsed at the firm-day level. The performance measures are in %-points. We include as controls the eigenvalue centrality of dealers ("Dealer Eig. Centrality") as in Maggio, Franzoni, Kermani, and Sommavilla (2019), the natural logarithm of the pound trade volume of dealers ("DealerVolume") and subsidiaries ("SubsidVolume") the natural logarithm of the number of daily transactions of dealers ("DealerTran.") and subsidiaries ("SubsidTran."). To reduce noise, we winsorise the sample at the 1%-level. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

8.6 Aggregate Connections and the Yield Curve

Table 17: Daily Changes in Yields and Aggregate Connections

	$\left\|\Delta Yield_t^{5Y}\right\|$			
	(1)	(2)	(3)	(4)
$\Delta \log (Connections_t)$	0.0277***		0.0233***	0.0234**
	(7.27)		(4.17)	(2.26)
$\Delta \log (Volume_t)$		0.0100***	0.0025	0.0025
		(6.11)	(1.07)	(1.06)
$\Delta \log (NumOfClients_t)$				-0.0001
				(-0.01)
N	1449	1449	1449	1449
R^2	0.040	0.030	0.041	0.041
	$\left\|\Delta\left(Yield_t^{25Y} - Yield_t^{1Y}\right)\right\|$			
	(1)	(2)	(3)	(4)
$\Delta \log (Connections_t)$	0.0209***		0.0173***	0.0236***
	(5.04)		(3.07)	(2.65)
$\Delta \log (Volume_t)$		0.0076***	0.0020	0.0019
		(4.84)	(1.00)	(0.92)
$\Delta \log (NumOfClients_t)$				-0.0082
				(-0.98)
N	1449	1449	1449	1449
R^2	0.072	0.064	0.072	0.073

Notes: this table regresses the absolute value of daily changes in the 5-year yield (Top Panel) and in the term spread (Bottom Panel) on daily changes in the logarithm of the total number of aggregate connections, the total number of clients and the total trading volume. The term spread is measured as the difference between the 25-year yield and the 1-year yield. The transaction-level data is collapsed at the day level yielding 1450 trading days spanning the period 4 Oct 2011 to 30 June 2017. Data on yields are from the Bank of England Financial Database. T-statistics, based on robust standard errors, are in parentheses. The coefficients for the deterministic variables (constant, linear and quadratic time trends) are not shown. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 18: Daily Changes in Price Dispersion and Aggregate Connections

	$\Delta \log(\Omega_t)$ – Weighted			
	(1)	(2)	(3)	(4)
$\Delta \log(Connections_t)$	1.6892***		1.2892***	1.2109***
	(10.96)		(5.42)	(2.96)
$\Delta \log(Volume_t)$		0.6446***	0.2312**	0.2330**
		(9.73)	(2.27)	(2.28)
$\Delta \log(NumOfClients_t)$				0.1028
				(0.24)
N	1449	1449	1449	1449
R^2	0.079	0.062	0.082	0.082

	$\Delta \log(\Omega_t)$ – Unweighted			
	(1)	(2)	(3)	(4)
$\Delta \log(Connections_t)$	1.5167***		1.1224***	0.9860**
	(10.60)		(5.04)	(2.51)
$\Delta \log(Volume_t)$		0.5879***	0.2280**	0.2310**
		(9.02)	(2.31)	(2.34)
$\Delta \log(NumOfClients_t)$				0.1790
				(0.42)
N	1449	1449	1449	1449
R^2	0.066	0.053	0.069	0.069

Notes: this table regresses the %-change in daily price dispersion weighted by trade size (6.3) (Top Panel) and unweighted (Bottom Panel) on daily changes in the logarithm of the total number of aggregate connections, the total market volume and the total number of clients. The transaction-level data is collapsed at the day level yielding 1450 trading days spanning the period 4 Oct 2011 to 30 June 2017. T-statistics, based on robust standard errors, are in parentheses. The coefficients for the deterministic variables (constant, linear and quadratic time trends) are not shown. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Appendix for Online Publication

A A 2-by-2 Model of Multi-dealer Trading

Consider many days of trading indexed by t and informed clients indexed by i. The fundamental value of the asset is random walk with a drift, formalized as $V_t = V_{t-1} + \varepsilon_t$, where ε_t is either 0 or 1 with equal probability, and independent across days. The innovation, ε_t, becomes public information at the end of the day and all positions are liquidated at the closing price V_t. During the day informed clients with a signal on ε_t, noise traders, and market makers trade determining the mid-day transaction prices P_{it}. The objective of clients is to maximize their trading profit $x_{it}(V_t - P_{it})$ each day by choosing to buy or sell one unit at the prevailing ask or bid prices respectively, i.e., choosing $x_{it} = \{1, -1\}$. The trading protocol is a modified version of Glosten and Milgrom (1985). Clients and noise traders seek bid and ask quotes from one or more risk neutral, market maker (MM) in each period. Just as in Glosten and Milgrom (1985), we assume that MMs are competititve, hence, their quotes are determined by a zero expected profit condition. Sampling quotes from more market makers might be costly.

To convey the intuition we focus on the simplest possible case. We consider four potential MMs, $m = R^i, N^i$ serving two clients $i = 1, 2$. Client i is assigned to MMs R^i, N^i (for regular and potential new comer). We assume that client i's signal, $s_t = B, S$ (for buy and sell) is informative:

$$\Pr(\varepsilon_t = 1 | s_t = B) = \frac{1}{2} + \Delta_{ti},$$

where $\Delta_{ti} > 0$ might vary across clients and time. Δ_{ti} is observable to clients and has the two point support of $\{\Delta_L, \Delta_H\}$ with $\Delta_H > \Delta_L$. Before a client observes her signal, she commits to a quote request function $\rho(\Delta_{ti}) : \{\Delta_L, \Delta_H\} \to \{R^i, (R^i, N^i)\}$ which describes the states when dealer i requests quotes from one or both dealers. The cost of the earlier is normalized to 0, while requesting two sets of quotes costs c. We think of c as a non-observable, non-pecuniary cost. It is a reduced form treatment to capture a search cost, or the reputational cost coming from the unmodelled future punishment from the dealer who provided a quote but did not received the trade. Importantly, client i is present in the market at t with only probability $(1 - \alpha)$. Even if the client requests two sets of bid and ask prices, eventually, she can trade only with one of the dealers. After observing the bids and asks she decides whether to buy or sell at one of those prices. Whenever client i is not requesting quotes from a given MM assigned to her, regardless it is by choice or because she is not present at that period, a noise trader requests quotes instead and buys or sells with equal probability. Therefore, MMs receive exactly one request for quotes in any given period, but might or might not trade. We assume that MMs observe Δ_{it} of their assigned client, but they do not observe whether she is present at the given period. That is, they cannot observe whether a quote request comes from client i or a noise trader. After trading, positions are liquidated at the realised true value V_t.

The following Proposition characterizes the equilibrium in this stage game. The client requests two quotes if and only if her information is sufficiently precise. In that case, she gets identical quotes and trades with each of the MMs with equal probability. The intuition relies on a simple idea. For fixed parameters, when the client i asks a quote only from R^i, R^i trades with an informed dealer with probability $(1-\alpha)$, while N^i trades with only noise traders. Therefore, the bid-ask spread provided by R^i is relatively wide, while the bid-ask spread provided by N^i is zero. When i asks a quote from both MMs and trades with only one of them randomly, R^i faces with an informed dealer with a probability $(1-\alpha)\pi$ only, where π is the mixing probability. Therefore, she will give better quotes to the client. In equilibrium, π has to adjust in a way that N^i wants to give identical quotes to R^i. Therefore, mixing between two dealers helps the client better hide her information, implying more favourable transaction prices. At the same time asking for two sets of quotes is costly. Hence, the client chooses to do so if and only if Δ_{ti} is sufficiently high.

Proposition 1 *Let*

$$\bar{\Delta} = \frac{1+\alpha}{\alpha(1-\alpha)}c$$

be within the support of Δ_{ti}.

1. *If* $\Delta_{ti} < \bar{\Delta}$, *the informed trader* i *trades only with* R^i *and the equilibrium bid ask quotes are*

$$A_t^{R^i}\left(\Delta_{ti} < \bar{\Delta}\right) = V_{t-1} + \frac{1}{2} + \Delta_{ti}(1-\alpha)$$
$$B_t^{R^i}\left(\Delta_{ti} < \bar{\Delta}\right) = V_{t-1} + \frac{1}{2} - \Delta_{ti}(1-\alpha)$$
$$A_t^{N^i}\left(\Delta_{ti} < \bar{\Delta}\right) = B^N\left(\Delta_{ti} < \bar{\Delta}\right) = V_{t-1} + \frac{1}{2}.$$

2. *If* $\Delta_{ti} > \bar{\Delta}$, *the informed trader* i *seeks quotes from both MM and trades with each*

with equal probability. The equilibrium bid ask quotes are

$$A_t^{R^i}\left(\Delta_{ti}>\bar{\Delta}\right)=A_t^{N^i}\left(\Delta_{ti}>\bar{\Delta}\right)=V_{t-1}+\frac{1}{2}+\Delta_{ti}\frac{1-\alpha}{1+\alpha}$$
$$B_t^{iR}\left(\Delta_{ti}>\bar{\Delta}\right)=B_t^{N^i}\left(\Delta_{ti}>\bar{\Delta}\right)=V_{t-1}+\frac{1}{2}-\Delta_{ti}\frac{1-\alpha}{1+\alpha}.$$

Proof. The quotes are derived by Bayes Rule. For example, the ask price provided by R^i to i when the MM understands that i trades with her with probability $\pi=1$ is

$$\begin{aligned}
&E\left(V_{t-1}+\varepsilon_t|\text{ observing a buy in } t,\pi=1\right)=\\
&\qquad=V_{t-1}+\frac{\left(\alpha\frac{1}{2}+(1-\alpha)\left(\frac{1}{2}+\Delta_{ti}\right)\right)\frac{1}{2}}{\left(\alpha\frac{1}{2}+(1-\alpha)\left(\frac{1}{2}+\Delta_{ti}\right)\right)\frac{1}{2}+\left(\alpha\frac{1}{2}+(1-\alpha)\left(1-\left(\frac{1}{2}+\Delta_{ti}\right)\right)\right)\frac{1}{2}}\\
&\qquad=V_{t-1}+\frac{1}{2}+\Delta_{ti}\left(1-\alpha\right).
\end{aligned}$$

When the trader observes quotes from both MMs, in equilibrium the two MMs have to submit the same quotes given the mixed strategy of acceptance from the trader. For this, the informed trader has to mix with probability half. In this case the ask price is given by

$$\begin{aligned}
&E\left(V_{t-1}+\varepsilon_t|\text{ observing a buy in } t,\pi=\frac{1}{2}\right)=\\
&\quad=V_{t-1}+\frac{\left(\alpha\frac{1}{2}+(1-\alpha)\left(\frac{1}{2}+\Delta_{ti}\right)\frac{1}{2}\right)\frac{1}{2}}{\left(\alpha\frac{1}{2}+(1-\alpha)\left(\frac{1}{2}+\Delta_{ti}\Delta\right)\frac{1}{2}\right)\frac{1}{2}+\left(\alpha\frac{1}{2}+(1-\alpha)\left(1-\left(\frac{1}{2}+\Delta_{ti}\right)\right)\frac{1}{2}\right)\frac{1}{2}}\\
&\quad=V_{t-1}+\frac{1}{2}+\Delta_{ti}\frac{1-\alpha}{1+\alpha}.
\end{aligned}$$

For a fixed Δ_{ti}, the expected benefit of transacting at more favourable prices implied by mixing, $\pi=\frac{1}{2}$, is

$$\begin{aligned}
\Sigma_{V_t=0,1}\Pr\left(\varepsilon_t|s=H,\Delta_{it}\right)&\left(\left(\varepsilon_t-\left(\frac{1}{2}+\Delta_{it}\frac{1-\alpha}{1+\alpha}\right)\right)-\left(\varepsilon_t-\left(\frac{1}{2}-\Delta_{it}\left(1-\alpha\right)\right)\right)\right)=\\
&=\Sigma_{V_t=0,1}\Pr\left(\varepsilon_t|s=H,\Delta_{it}\right)\left(\Delta_{it}\left(1-\alpha\right)-\Delta_{it}\frac{1-\alpha}{1+\alpha}\right)\\
&=\Delta_{it}\alpha\frac{1-\alpha}{1+\alpha},
\end{aligned}$$

which is increasing in Δ_{ti}. The indifference condition given cost c determines $\bar{\Delta}$. ■

To picture the implied time-series and cross-sectional evolution of prices and trades, we assume that for each client i, this subgame is repeated in many time-periods. These games are independent from each other because all random variables are redrawn in each period and because the MMs are disjunct across the two clients. Suppose that $\Delta_H > \bar{\Delta} > \Delta_L$. The correlation structure across time and clients in Δ_{ti} can be arbitrary.

To see the implications, it is useful to compare implied profits of clients in the two states. (Because of symmetry, it is sufficient to calculate the implied expected profit conditional on a signal $s = B$.)

$$\begin{aligned}
\Pi(\Delta_H) &= \Sigma_{V_t=0,1} \Pr(\varepsilon_t | s = H, \Delta_H) \left(\varepsilon_t - \left(\frac{1}{2} + \Delta_H \frac{1-\alpha}{1+\alpha}\right)\right) \\
&= \left(\frac{1}{2} + \Delta_H\right)(1 - \left(\frac{1}{2} + \Delta_H \frac{1-\alpha}{1+\alpha}\right)) + \left(\frac{1}{2} - \Delta_H\right)\left(0 - \left(\frac{1}{2} + \Delta_H \frac{1-\alpha}{1+\alpha}\right)\right) \\
&= 2\alpha \frac{\Delta_H}{1+\alpha}.
\end{aligned}$$

$$\begin{aligned}
\Pi(\Delta_L) &= \Sigma_{V_t=0,1} \Pr(\varepsilon_t | s = H, \Delta_L) \left(\varepsilon_t - \left(\frac{1}{2} - \Delta_L (1-\alpha)\right)\right) \\
&= \left(\frac{1}{2} + \Delta_L\right)(1 - \left(\frac{1}{2} + \Delta_L (1-\alpha)\right)) + \left(\frac{1}{2} - \Delta_L\right)\left(0 - \left(\frac{1}{2} + \Delta_L (1-\alpha)\right)\right) \\
&= \alpha \Delta_L.
\end{aligned}$$

Clearly,

$$\Pi(\Delta_H) - \Pi(\Delta_L) = \alpha \left(\frac{1-\alpha}{1+\alpha} \Delta_H + (\Delta_H - \Delta_L)\right) > 0,$$

for any parameter values. Note also, that $\Pi(\Delta_H) - \Pi(\Delta_L)$ is increasing in Δ_H and in $(\Delta_H - \Delta_L)$.

Consider an interval with multiple, say, D periods. Within this interval, in each period when $\Delta_{it} = \Delta_L$, client i trades with only R^i with probability 1. In each period when $\Delta_{it} = \Delta_H$, with probability $\frac{1}{2}$ she trades with N^i. Hence, if ξ_{iD} is a counting process for the periods with $\Delta_{it} = \Delta_H$ within D, then the expected number of connections of i within interval D is $1 + \frac{\xi_{iD}}{2}$, an increasing function of ξ_{iD}. That is, the number of connections within an interval is a proxy for the number of periods with an interval where the information precision of the client is Δ_H. These observations give the following hypotheses.

Hypothesis 1 *More connections for a client i in a given interval should be associated with higher trading profit.*

Hypothesis 2 *The positive relationship between connections for a client i in a given interval and trading profit should be stronger in periods with more precise private information.*

For later use, note that aggregating connections over clients in a given interval D is a proxy for the total private information present in the market.

Next, observe that the difference in the trading profit across $\Delta_{it} = \Delta_H$ and $\Delta_{it} = \Delta_L$ comes from two sources: the change in probability that the client trades the right direction, and the change in the transaction price. On one hand, the probability that client i is trading at the right direction, buying before the price moves up and selling before the price moves down, is $\frac{1}{2} + \Delta_{ti}$, an increasing function of Δ_{ti}. This gives the second hypothesis, which we refer to in the text as the anticipation component.

Hypothesis 3 *More connections for a client i in a given interval should be associated with a stronger connection between her buy (sell) transactions and subsequent positive (negative) returns.*

On the other hand, the transaction price can be more or less favorable when $\Delta_{it} = \Delta_H$. In particular, if and only if

$$\frac{\Delta_H}{\Delta_L} < 1 + \alpha,$$

a client buys (sells) at a lower (higher) price when she is more informed. This relationship drives the sign of the relationship between connections and the transaction component defined in the text. The reason for the ambiguous result is that the client's ability to hide her higher quality signal better by trading with more dealers is limited. When

$$1 + \alpha < \frac{\Delta_H}{\Delta_L}, \tag{A.1}$$

then a client with higher information precision mixing between the two dealers gets a less favourable price than the client with the lower information precision who trades with one dealer only. Of course, the high precision client's price is still more favourable than the price which would prevail if she were to trade with her regular dealer only. Otherwise, there would be a profitable deviation in equilibrium.

Finally, we turn to price discovery. For this, we calculate the expected average transaction price when the innovation is $\varepsilon_t = 1$ in each possible scenarios with respect to the type of traders arriving. First, when both traders have high precision signals, both request quotes from both of their assigned MMs, but they trade only at one of those quotes. The average transaction price is

$$
\begin{aligned}
&E\left(\frac{P_{1t}+P_{2t}}{2}|\varepsilon_t=1,\Delta_H,\Delta_H\right)=\\
=&V_{t-1}+\left(\frac{1}{2}+\Delta_H\right)\left(\frac{1}{2}+\Delta_H\right)\left(\left(\frac{1}{2}+\Delta_H\frac{1-\alpha}{1+\alpha}\right)\right)\\
&+\left(\frac{1}{2}-\Delta_H\right)\left(\frac{1}{2}-\Delta_H\right)\left(\left(\frac{1}{2}-\Delta_H\frac{1-\alpha}{1+\alpha}\right)\right)\\
&+2\left(\frac{1}{2}+\Delta_H\right)\left(\frac{1}{2}-\Delta_H\right)\left(\frac{1}{2}\right)\\
=&V_{t-1}+\frac{1}{2}+2\Delta_H^2\frac{1-\alpha}{1+\alpha}.
\end{aligned}
$$

If both clients have low precision signals, each request quotes only from R^i and N^i trades with liquidity traders at price $V_{t-1}+\frac{1}{2}$. That is, the average transaction price is

$$
\begin{aligned}
&E\left(\frac{P_{1t}\left(R^1\right)+P_{2t}\left(R^2\right)+P_{1t}\left(N^1\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\Delta_L,\Delta_L\right)=\\
=&V_{t-1}+\left(\frac{1}{2}+\Delta_L\right)\left(\frac{1}{2}+\Delta_L\right)\left(\frac{1}{2}\left(\frac{1}{2}+\Delta_L\left(1-\alpha\right)\right)+\frac{1}{2}\frac{1}{2}\right)\\
&+\left(\frac{1}{2}-\Delta_L\right)\left(\frac{1}{2}-\Delta_L\right)\left(\frac{1}{2}\left(\frac{1}{2}-\Delta_L\left(1-\alpha\right)\right)+\frac{1}{2}\frac{1}{2}\right)\\
&+2\left(\frac{1}{2}+\Delta_L\right)\left(\frac{1}{2}-\Delta_L\right)\left(\frac{1}{2}\right)\\
=&V_{t-1}+\frac{1}{2}+\Delta_L^2\left(1-\alpha\right).
\end{aligned}
$$

When the first client has low precision, while the second one has high precision, the average price is

$$
\begin{aligned}
&E\left(\frac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{1t}}{3}|\varepsilon_t=1,\Delta_L,\Delta_H\right)=\\
=&V_{t-1}+\left(\frac{1}{2}+\Delta_L\right)\left(\frac{1}{2}+\Delta_H\right)\left(\left(\frac{1}{2}+\frac{1}{3}\left(\Delta_L\left(1-\alpha\right)+\Delta_H\frac{1-\alpha}{1+\alpha}\right)\right)\right)\\
&+\left(\frac{1}{2}-\Delta_H\right)\left(\frac{1}{2}-\Delta_L\right)\left(\left(\frac{1}{2}-\frac{1}{3}\left(\Delta_L\left(1-\alpha\right)+\Delta_H\frac{1-\alpha}{1+\alpha}\right)\right)\right)\\
&+\left(\frac{1}{2}+\Delta_H\right)\left(\frac{1}{2}-\Delta_L\right)\left(\frac{1}{2}+\frac{1}{3}\left(\Delta_H\frac{\left(1-\alpha\right)}{1+\alpha}-\Delta_L\left(1-\alpha\right)\right)\right)\\
&+\left(\frac{1}{2}-\Delta_H\right)\left(\frac{1}{2}+\Delta_L\right)\left(\frac{1}{2}-\frac{1}{3}\left(\Delta_H\frac{\left(1-\alpha\right)}{1+\alpha}-\Delta_L\left(1-\alpha\right)\right)\right)\\
=&V_{t-1}+\frac{1}{2}+\frac{2}{3}\left(1-\alpha\right)\left(\frac{\Delta_H^2}{1+\alpha}+\Delta_L^2\right).
\end{aligned}
$$

With similar calculations we get the expected average price for the case when at least one of the clients are not present (which we deonte by $\oslash$), hence replaced by two noise traders:

$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}\left(R^2\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\oslash,\oslash\right)=V_{t-1}+\frac{1}{2}$$
$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}\left(R^2\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\oslash,\Delta_L\right)=V_{t-1}+\frac{1}{2}+\frac{1}{2}\Delta_L^2\left(1-\alpha\right)$$
$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}}{3}|\varepsilon_t=1,\oslash,\Delta_H\right)=V_{t-1}+\frac{1}{2}+\frac{1}{2}\Delta_H^2\frac{1-\alpha}{1+\alpha}.$$

It is easy to check that

$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}\left(R^2\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\oslash,\Delta_L\right)<$$
$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}}{3}|\varepsilon_t=1,\oslash,\Delta_H\right),E\left(\tfrac{P_{1t}\left(R^1\right)+P_{2t}\left(R^2\right)+P_{1t}\left(N^1\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\Delta_L,\Delta_L\right)$$

and

$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}}{3}|\varepsilon_t=1,\oslash,\Delta_H\right),E\left(\tfrac{P_{1t}\left(R^1\right)+P_{2t}\left(R^2\right)+P_{1t}\left(N^1\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\Delta_L,\Delta_L\right)<$$
$$E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{1t}}{3}|\varepsilon_t=1,\Delta_L,\Delta_H\right)<E\left(\tfrac{P_{1t}+P_{2t}}{2}|\varepsilon_t=1,\Delta_H,\Delta_H\right).$$

Recall that aggregate connections $\Sigma_i\xi_{iD}$ is increasing in the fraction of high precision clients in the market. Therefore, with the caveat that the comparison of $E\left(\tfrac{P_{1t}\left(R^1\right)+P_{1t}\left(N^1\right)+P_{2t}}{3}|\varepsilon_t=1,\oslash,\Delta_H\right)$ and $E\left(\tfrac{P_{1t}\left(R^1\right)+P_{2t}\left(R^2\right)+P_{1t}\left(N^1\right)+P_{2t}\left(N^2\right)}{4}|\varepsilon_t=1,\Delta_L,\Delta_L\right)$ depends on the parameters, we form the following hypothesis.

Hypothesis 4 *Periods with higher aggregate connections should be associated with larger absolute innovations in prices.*

B Additional Tables and Figures

B.1 Sensitivity Analysis of the Baseline Regression

Table 19: Client Connections and Trading Performance: Excluding Client Fixed Effects

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client Connections	0.0016***	0.0024***	0.0021**	0.0022**	0.0021**
	(2.70)	(2.97)	(2.49)	(2.40)	(2.15)
Volume	0.0027**	0.0026**	0.0023*	0.0026**	0.0028*
	(2.53)	(2.07)	(1.73)	(2.00)	(1.90)
Transactions	-0.0030	-0.0047*	-0.0038	-0.0059**	-0.0050
	(-1.62)	(-1.95)	(-1.42)	(-2.01)	(-1.48)
N	20840	20840	20840	20840	20840
R^2	0.006	0.006	0.006	0.007	0.007
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	No	No	No	No

(a) Turnover-weighted Trading Performance

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client Connections	0.0027***	0.0033***	0.0033***	0.0035***	0.0038***
	(4.50)	(4.61)	(3.88)	(4.08)	(4.04)
Volume	0.0043***	0.0045***	0.0040***	0.0041***	0.0039**
	(4.14)	(3.64)	(2.85)	(2.95)	(2.46)
Transactions	-0.0079***	-0.0101***	-0.0094***	-0.0112***	-0.0108***
	(-4.11)	(-4.46)	(-3.72)	(-4.03)	(-3.46)
N	20840	20840	20840	20840	20840
R^2	0.010	0.010	0.008	0.009	0.008
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	No	No	No	No

(b) Unweighted Trading Performance

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on our connectivity measures (4.2). The transaction-level data is collapsed at the client-month level. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client and the natural logarithm of the number of monthly transactions. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 20: Client Connections and Trading Performance: Including Public Clients

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.0015	0.0029**	0.0033**	0.0040**	0.0039**
Connections	(1.62)	(2.55)	(2.49)	(2.44)	(2.38)
Volume	-0.0007	-0.0001	-0.0001	-0.0017	-0.0022
	(-0.38)	(-0.06)	(-0.03)	(-0.57)	(-0.62)
Transactions	0.0042	0.0018	-0.0007	-0.0027	-0.0017
	(1.58)	(0.46)	(-0.17)	(-0.55)	(-0.28)
N	22843	22843	22843	22843	22843
R^2	0.038	0.039	0.036	0.034	0.035
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(a) Turnover-weighted Trading Performance

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.0016*	0.0029***	0.0029**	0.0039**	0.0041**
Connections	(1.83)	(2.76)	(2.14)	(2.55)	(2.58)
Volume	-0.0013	0.0009	0.0025	0.0014	0.0001
	(-0.75)	(0.43)	(0.98)	(0.52)	(0.04)
Transactions	0.0019	-0.0023	-0.0043	-0.0051	-0.0029
	(0.62)	(-0.60)	(-1.03)	(-1.11)	(-0.56)
N	22843	22843	22843	22843	22843
R^2	0.051	0.047	0.044	0.039	0.038
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Unweighted Trading Performance

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on our connectivity measures (4.2). The transaction-level data is collapsed at the client-month level. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client and the natural logarithm of the number of monthly transactions. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 21: Client Connections and Trading Performance: Using Eigenvalue-Centrality

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.3546*	0.7021***	0.8340***	0.9727***	1.0242***
Centrality	(1.74)	(2.77)	(2.87)	(2.80)	(2.79)
Volume	-0.0005	0.0003	0.0002	-0.0014	-0.0019
	(-0.24)	(0.14)	(0.08)	(-0.45)	(-0.52)
Transactions	0.0033	0.0003	-0.0025	-0.0049	-0.0040
	(1.18)	(0.08)	(-0.58)	(-0.96)	(-0.64)
N	20839	20839	20839	20839	20839
R^2	0.037	0.039	0.036	0.034	0.036
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(a) Turnover-weighted Trading Performance

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.3845**	0.6661***	0.6706**	0.8549**	0.9978***
Centrality	(2.05)	(2.79)	(2.19)	(2.54)	(2.85)
Volume	-0.0011	0.0014	0.0028	0.0019	0.0006
	(-0.66)	(0.61)	(1.04)	(0.69)	(0.19)
Transactions	0.0011	-0.0038	-0.0059	-0.0070	-0.0052
	(0.37)	(-0.96)	(-1.34)	(-1.45)	(-0.98)
N	20839	20839	20839	20839	20839
R^2	0.051	0.048	0.044	0.039	0.039
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Unweighted Trading Performance

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on our connectivity measures (4.2). The transaction-level data is collapsed at the client-month level. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client and the natural logarithm of the number of monthly transactions. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 22: Client Connections at Daily Frequency

	Mean	Median	St.dev	Within St.dev	N
Connections	3.19	3	2.33	1.46	103,199

Note: the table presents summary statistics on client connections, defined as the number of dealers a client trades with on a given trading day. "Within St.dev" is the standard deviation of the estimated residual $\varepsilon_{i,t}$ from the regression $Connections_{i,t} = \alpha_i + \mu_t + \varepsilon_{i,t}$. The sample is based on trading days on which a client has more than two transactions.

B.2 Daily Data

B.2.1 Performance Regressions

Table 23: Client Connections and Trading Performance Using Daily Data

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.0005	0.0009	0.0028**	0.0043***	0.0028*
Connections	(0.65)	(1.01)	(2.54)	(3.23)	(1.92)
Volume	0.0012	0.0015	0.0018	0.0015	0.0023
	(1.03)	(1.13)	(1.19)	(0.89)	(1.22)
Tran.	-0.0006	-0.0025	-0.0086***	-0.0081**	-0.0093**
	(-0.27)	(-0.90)	(-2.62)	(-2.10)	(-2.29)
N	103565	103565	103565	103565	103565
R^2	0.035	0.035	0.034	0.033	0.033
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

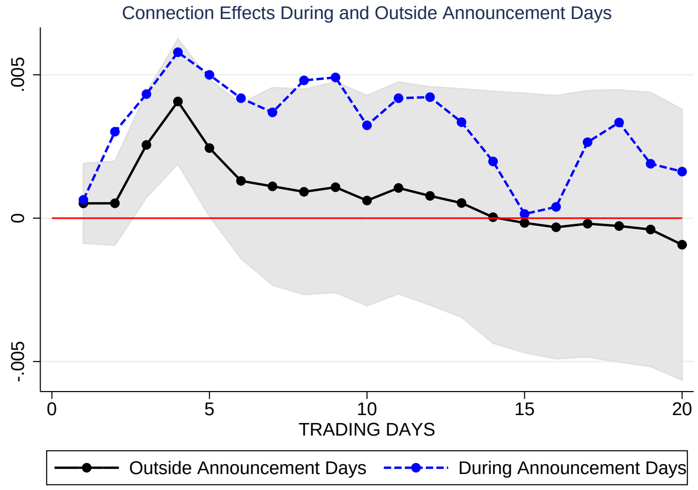
(a) Turnover-weighted Trading Performance

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
Client	0.0012	0.0015	0.0027**	0.0042***	0.0030*
Connections	(1.47)	(1.57)	(2.13)	(2.63)	(1.90)
Volume	-0.0004	0.0004	0.0010	0.0014	0.0014
	(-0.44)	(0.32)	(0.73)	(0.89)	(0.78)
Tran.	0.0007	-0.0011	-0.0060	-0.0079*	-0.0077
	(0.32)	(-0.36)	(-1.56)	(-1.74)	(-1.58)
N	103565	103565	103565	103565	103565
R^2	0.039	0.037	0.036	0.036	0.036
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Unweighted Trading Performance

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on our connectivity measures (4.2). The transaction-level data is collapsed at the client-day level. The performance measures are in %-points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of daily transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use day-client observations that are based on at least 2 transactions in the day. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Figure 7: Daily Performance Regressions over 0-20 day Horizons



Notes: Panel a plots the estimated β coefficients from our baseline regression 4.2 up to 20-day horizon ($T = 20$), using the value weighted performance variable as the regressand, measured in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of daily transactions ("Transactions"). Panel b plots the results after including a dummy (interacted with connectedness) indicating the trading days that coincided with MPC announcements and release of inflation data. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least three daily transactions. The shaded area denotes the 90% confidence band, It is based on robust standard errors, using two-way clustering at the day and the client level.

B.2.2 Connectivity and Predicting the Order Flow

Table 24: Weighted Trading Performance on Trading Days with High Covariance with the **Total Market** Order Flow

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
$Q^{Total}_{i,t} = 1$	0.0181***	0.0255***	0.0284***	0.0257***	0.0231***
	(4.75)	(5.24)	(4.73)	(3.72)	(3.00)
Volume	0.0018	0.0021	0.0025*	0.0023	0.0032*
	(1.50)	(1.58)	(1.71)	(1.41)	(1.68)
Tran.	-0.0008	-0.0024	-0.0063**	-0.0039	-0.0067*
	(-0.39)	(-0.92)	(-2.16)	(-1.14)	(-1.95)
N	105190	105190	105190	105190	105190
R^2	0.037	0.036	0.036	0.034	0.033
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(a) Covariance with 1-day Ahead

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
$Q^{Total}_{i,t} = 1$	-0.0013	0.0098**	0.0214***	0.0229***	0.0268***
	(-0.37)	(2.09)	(3.64)	(3.34)	(3.55)
Volume	0.0019	0.0021	0.0025*	0.0023	0.0031
	(1.53)	(1.58)	(1.66)	(1.37)	(1.63)
Tran.	-0.0006	-0.0022	-0.0062**	-0.0038	-0.0066*
	(-0.29)	(-0.84)	(-2.10)	(-1.11)	(-1.91)
N	105041	105041	105041	105041	105041
R^2	0.035	0.035	0.035	0.033	0.033
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Covariance with 3-day Ahead

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on a dummy $Q^{Total}_{i,t}$ that takes value 1 if on day t the order flow of client i has a covariance (see measure 4.2) with the future order flow of the market that is higher than the median (based on all trading days of the given client).. The transaction-level data is collapsed at the client-day level. The performance measures are in %-points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of daily transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use day-client observations that are based on at least 2 transactions in the day. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 25: Weighted Trading Performance on Trading Days with High Covariance with the Market Order Flow Intermediated by **Own Dealers**

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
$Q_{i,t}^{Own} = 1$	0.0067***	0.0081**	0.0094**	0.0100**	0.0118***
	(2.89)	(2.49)	(2.38)	(2.31)	(2.61)
Volume	0.0013	0.0016	0.0021	0.0020	0.0027
	(1.11)	(1.23)	(1.43)	(1.21)	(1.42)
Tran.	-0.0002	-0.0017	-0.0057*	-0.0035	-0.0064*
	(-0.08)	(-0.64)	(-1.93)	(-1.00)	(-1.83)
N	103103	103103	103103	103103	103103
R^2	0.035	0.035	0.034	0.033	0.033
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(a) Covariance with 1-day Ahead

	(1) 1-day	(2) 2-day	(3) 3-day	(4) 4-day	(5) 5-day
$Q_{i,t}^{Own} = 1$	-0.0017	0.0021	0.0066*	0.0104**	0.0105**
	(-0.80)	(0.68)	(1.76)	(2.50)	(2.36)
Volume	0.0013	0.0016	0.0021	0.0020	0.0027
	(1.09)	(1.21)	(1.40)	(1.19)	(1.40)
Tran.	-0.0000	-0.0015	-0.0056*	-0.0034	-0.0063*
	(-0.00)	(-0.58)	(-1.88)	(-0.98)	(-1.78)
N	102958	102958	102958	102958	102958
R^2	0.035	0.035	0.034	0.033	0.033
Time FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes

(b) Covariance with 3-day Ahead

Notes: this table regresses the value-weighted (panel A) and unweighted (panel B) trading performance at different time horizons on a dummy $Q_{i,t}^{Own}$ that takes value 1 if on day t the order flow of client i has a covariance (see measure 4.2) with the future order flow of its own dealers that is higher than the median (based on all trading days of the given client). Own dealers ($g = Own$) are the ones that the client traded with on the day of the trade (for which the trading performance is calculated) as well as during the past 10 trading days. The transaction-level data is collapsed at the client-day level. The performance measure is in %-points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of daily transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use day-client observations that are based on at least 2 transactions in the day. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 26: Client Connectivity and Covariance with the Aggregate Order Flow

	1-Day	3-Day	5-Day
Client	0.0050**	0.0006	0.0011
Connections	(2.40)	(0.36)	(0.48)
Volume	0.0040**	0.0025	0.0026
	(2.29)	(1.53)	(1.58)
Tran.	0.0134**	0.0138**	0.0108*
	(2.36)	(2.52)	(1.78)
N	103094	102949	102816
R^2	0.023	0.024	0.023
Time/Client FE	Yes/Yes	Yes/Yes	Yes/Yes

(a) Total Client Order Flow

	Traded with Dealers on Trading Day as well as During Past 10 Days			Traded with Dealers only on Trading Day		
	1-Day	3-Day	5-Day	1-Day	3-Day	5-Day
Client	0.0068***	0.0052***	0.0030*	-0.0007	-0.0008	-0.0007
Connections	(2.87)	(3.67)	(1.87)	(-0.79)	(-0.77)	(-0.69)
Volume	0.0038**	0.0014	0.0030**	-0.0006	-0.0002	-0.0003
	(2.55)	(1.16)	(2.34)	(-0.77)	(-0.16)	(-0.36)
Tran.	0.0092*	0.0038	0.0020	0.0030*	0.0031	-0.0002
	(1.81)	(0.94)	(0.46)	(1.76)	(1.51)	(-0.11)
N	103103	102958	102825	103103	102958	102825
R^2	0.023	0.021	0.022	0.021	0.021	0.022
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes	Yes

(b) Total Client Order Flow via Own Dealers

Notes: this table regresses different versions of the covariance measure 5.3 on our connectivity measure and controls (5.5). The transaction-level data is collapsed at the client-day level. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of day transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use day-client observations that are based on at least 2 transactions on the day. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

B.3 Announcements Including News About US FOMC and UK Real Variables

Table 27: Turnover-weighted Performance: Trading Days With and Without Inflation & Monetary Policy News; Including News About UK Real Variables and US Monetary Policy

	1-day	2-day	3-day	4-day	5-day
Connections	0.0022**	0.0027**	0.0027*	0.0037**	0.0032
	(2.17)	(2.05)	(1.81)	(2.12)	(1.61)
Connections$\times ANN'$	0.0007	0.0014**	0.0026***	0.0019**	0.0020
	(1.20)	(2.18)	(3.16)	(2.05)	(1.63)
Volume	0.0005	0.0027	0.0046*	0.0036	0.0022
	(0.30)	(1.22)	(1.76)	(1.26)	(0.72)
Tran.	0.0005	-0.0016	-0.0065	-0.0072	-0.0035
	(0.19)	(-0.42)	(-1.58)	(-1.65)	(-0.66)
N	37733	37733	37733	37733	37733
R^2	0.024	0.024	0.022	0.021	0.022
Client FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes

Notes: this table regresses the value-weighted trading performance at different time horizons on client connections and connections interacted with a dummy that takes value 1 when the trading days coincide macroeconomic announcements. The transaction-level data is collapsed at the client-month-dummy level, i.e. for each month and each client we compute two sets of performance measures and controls, one set based on announcements days the other set based on trading days without announcements. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 28: Turnover-weighted Performance: Trading Days With and Without Inflation & Monetary Policy News; Including News About UK Real Variables and US Monetary Policy

	1-day	2-day	3-day	4-day	5-day
$\beta_{ANN'}$	0.0007	0.0015**	0.0021**	0.0016*	0.0019
	(1.14)	(2.35)	(2.52)	(1.82)	(1.66)
N	33674	33674	33674	33674	33674
R^2	0.506	0.512	0.520	0.527	0.525
Client-Month FE	Yes	Yes	Yes	Yes	Yes

Notes: this table regresses the value-weighted trading performance at different time horizons on client connections and connections interacted with a dummy that takes value 1 when the trading days coincide macroeconomic announcements. The transaction-level data is collapsed at the client-month-dummy level, i.e. for each month and each client we compute two sets of performance measures, one set based on announcements days the other set based on trading days without announcements. The performance measures are in basis points. We include as a control the natural logarithm of the pound trade volume of each client ("Volume") and the natural logarithm of the number of monthly transactions ("Tran."). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

B.4 Aggregate Connections and the Yield Curve

Table 29: Daily Changes in Yields and Aggregate Connections

	$\left\|\Delta Yield_t^{10Y}\right\|$			
	(1)	(2)	(3)	(4)
$\Delta\log(Connections_t)$	0.0308***		0.0297***	0.0326***
	(6.70)		(4.33)	(2.73)
$\Delta\log(Volume_t)$		0.0102***	0.0006	0.0006
		(5.43)	(0.23)	(0.21)
$\Delta\log(NumOfClients_t)$				-0.0038
				(-0.32)
N	1449	1449	1449	1449
R^2	0.041	0.028	0.041	0.041
	$\left\|\Delta Yield_t^{25Y}\right\|$			
	(1)	(2)	(3)	(4)
$\Delta\log(Connections_t)$	0.0253***		0.0191***	0.0223**
	(6.10)		(3.16)	(2.15)
$\Delta\log(Volume_t)$		0.0097***	0.0036	0.0035
		(5.82)	(1.49)	(1.45)
$\Delta\log(NumOfClients_t)$				-0.0042
				(-0.41)
N	1449	1449	1449	1449
R^2	0.048	0.042	0.049	0.049

Notes: this table regresses the absolute value of daily changes in the 10-year yield (Top Panel) and in the 25-year yield (Bottom Panel) on daily changes in the logarithm of the total number of aggregate connections, the total number of clients and the total number of transactions. The transaction-level data is collapsed at the day level yielding 1450 trading days spanning the period 4 Oct 2011 to 30 June 2017. Data on yields are from the Bank of England Financial Database. T-statistics, based on robust standard errors, are in parentheses. The coefficients for the deterministic variables (constant, linear and quadratic time trends) are not shown. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

B.5 The Centrality of Dealers

Table 30: Clients' Trading Performance and the Centrality of Dealers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	0-day				3-day			
Eig. Centrality	0.1068**	0.1160***	0.0689	0.0667	0.2117**	0.2029*	0.0294	-0.0147
	(2.47)	(2.88)	(1.40)	(1.46)	(2.05)	(1.86)	(0.22)	(-0.10)
Dealer Volume			0.0037	0.0041			0.0172*	0.0200**
			(1.12)	(1.25)			(1.89)	(2.04)
Average			-0.0037	-0.0013			-0.0153*	-0.0155*
Trade Size			(-1.34)	(-0.45)			(-1.91)	(-1.74)
N	105564	101852	105564	101852	105564	101852	105564	101852
R^2	0.011	0.210	0.011	0.210	0.008	0.226	0.008	0.226
Time FE	Yes	No	Yes	No	Yes	No	Yes	No
Dealer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	No	Yes	No	Yes	No	Yes	No
Client-Time FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: this table regresses the value-weighted trading performance at 0-day (Columns 1-4) and 3-day (Columns 5-8) horizons on dealer centrality (Maggio, Franzoni, Kermani, and Sommavilla, 2019). The transaction-level data is collapsed at the client-dealer-month level. The performance measures are in %-points. We include as a control the natural logarithm of the pound trade volume of dealers ("Dealer Volume") and the natural logarithm of the average transaction size of dealers ("Average Trade Size"). To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

Table 31: Client Connections and Dealer Centrality

	0-day	1-day	2-day	3-day	4-day	5-day
Eig. Centrality	0.1068**	0.0903	0.1232	0.2117**	0.2161*	0.1275
	(2.47)	(1.23)	(1.51)	(2.05)	(1.78)	(0.86)
N	105564	105564	105564	105564	105564	105564
R^2	0.011	0.009	0.009	0.008	0.008	0.008
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Dealer FE	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes	Yes

(a) Clients' Trading Performance and the Centrality of Dealers

	0-day	1-day	2-day	3-day	4-day	5-day
Client	0.0008**	0.0016**	0.0020**	0.0014	0.0012	0.0016
Connections	(2.14)	(2.59)	(2.33)	(1.25)	(0.89)	(1.18)
N	105564	105564	105564	105564	105564	105564
R^2	0.011	0.009	0.009	0.008	0.007	0.008
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Dealer FE	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes	Yes

(b) Clients' Trading Performance and Client Connections

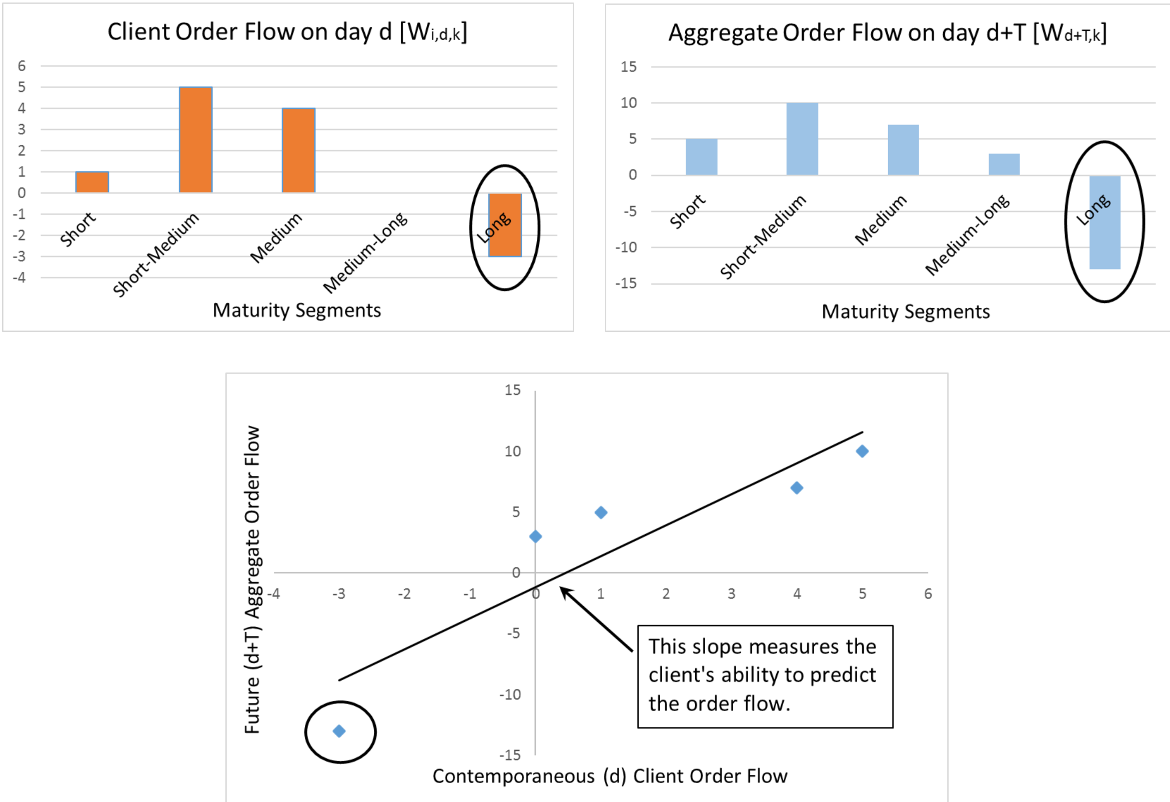
	0-day	1-day	2-day	3-day	4-day	5-day
Eig. Centrality	0.1093**	0.0955	0.1297	0.2163**	0.2200*	0.1327
	(2.53)	(1.30)	(1.59)	(2.09)	(1.81)	(0.89)
Client	0.0008**	0.0016**	0.0020**	0.0014	0.0012	0.0016
Connections	(2.20)	(2.62)	(2.36)	(1.28)	(0.93)	(1.20)
N	105564	105564	105564	105564	105564	105564
R^2	0.011	0.009	0.009	0.008	0.008	0.008
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Dealer FE	Yes	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	Yes	Yes	Yes	Yes

(c) Clients' Trading Performance: Including both the Centrality of Dealers and Client Connections

Notes: this table regresses the value-weighted trading performance at different time horizons on dealer centrality (Panel A), client connections (Panel B) and on both variables at the same time (Panel C). The transaction-level data is collapsed at the client-dealer-month level. The performance measures are in %-points. To reduce noise, we winsorise the sample at the 1%-level and use month-client observations that are based on at least 2 transactions in the month. T-statistics in parentheses are based on robust standard errors, using two-way clustering at the month and the client level. Asterisks denote significance levels (* p<0.1, ** p<0.05, *** p<0.01).

B.6 Illustrating Our Order Flow Covariance Measure

Figure 8: Illustrating Our Order Flow Covariance Measure



Notes: this figure illustrates how we measure the ability of a client to predict the order flow (equation 5.3). The top left panel shows for a hypothetical client i, on trading day d the net position in each k duration segments. The top right panel shows the accumulated net position of the market T days later in each k duration segments. The scatter plot in the bottom panel illustrates the cross-sectional nature of our measure.