Size Discount and Size Penalty: Trading Costs in Bond Markets

Gabor Pinter
Bank for International Settlements, Switzerland

Chaojun Wang
The Wharton School, University of Pennsylvania, USA

Junyuan Zou
INSEAD, France

We show that larger trades incur lower trading costs in government bond markets ("size discount"), but costs increase in trade size after controlling for client identity ("size penalty"). The size discount is driven by the cross-client variation of larger traders obtaining better prices, consistent with theories of trading with imperfect competition. The size penalty, driven by the within-client variation, is larger for corporate bonds, during major macroeconomic surprises and during COVID-19. These differences are larger among more sophisticated clients, consistent with information-based theories. (*JEL* G12, G14, G24)

Received XXXX XX, XXXX; editorial decision XXXX XX, XXXX by Editor XXXXXXXXXXXX.

It is well documented that larger trades incur lower trading costs ("size discount") in various over-the-counter (OTC) financial markets. The size discount is consistent with theories of bilateral trading with imperfect competition, which predict that larger trades get more favourable prices because dealers' bargaining power decreases in trader

We would like to thank Tarun Chordia, Bernard Dumas, Thierry Foucault, Vincent Glode, Itay Goldstein, Joel Hasbrouck, Edith Hotchkiss (discussant), John Kuong, Pete Kyle, Dan Li (discussant), Weiling Liu (discussant), Dmitry Livdan (discussant), Albert J. Menkveld, Paolo Pasquariello, Joel Peress, Tarun Ramadorai (the editor), Robert Stambaugh, Lucian Taylor, Xavier Vives, Alex Zhou, Haoxiang Zhu, and two anonymous referees for helpful comments. We also thank seminar participants at INSEAD, Wharton, the 2021 Stern/Salomon Microstructure Meeting, the 2021 EFA meeting, the 2022 AFA meeting, and the 2022 CICF meeting for valuable discussions. The empirical work was completed while Pinter was at the Bank of England. The views expressed in this paper are those of the authors, and not necessarily those of the Bank for International Settlements (BIS) or the Bank of England. Send correspondence to Gabor Pinter, gabor.pinter@bis.org.

Published by Oxford University Press on behalf of The Society for Financial Studies 2014.
doi:10.1093/rfs/Sample Advance Access publication September 21, 2014

size and larger traders tend to trade larger amounts.[1] However, theories of information asymmetry and inventory imbalances predict "size penalty," in that larger trades would be executed at less favourable prices because of dealers' fear of adverse selection[2] or higher inventory costs.[3]

We reconcile this tension by decomposing the cost-size relationship into cross-client and within-client variations, finding size discount in the cross-section and size penalty in the time series. Given the scant evidence on the size penalty in OTC markets (compared to ample evidence on the size discount), we analyze the drivers of the size penalty in further detail by applying difference-in-differences methods. Our analysis illustrates the effects of different market frictions on trading costs, and our findings point to an independent role of information-based theories, controlling for inventory- and liquidity-based explanations, in driving the size penalty.

Our paper exploits a nonanonymous trade-level data set to study the determinants of trading costs in bond markets, with our baseline sample covering trades in the U.K. government bond market over the period 2011-2017. The data set covers close to the universe of secondary market transactions, and importantly, it contains the *identities of both counterparties* for each transaction. Therefore, unlike other data sets (e.g., TRACE) typically used in the literature, our data set allows one to distinguish between client-specific characteristics (such as traders' size and type) and transaction-specific characteristics (such as trade size) in determining trading costs. We also identify clients who actively trade in U.K. government bonds as well as in U.K. corporate bonds, which allows us to compare the cost-size relation not just across clients but also *across markets.*

Our empirical analysis yields six main results. First, larger trades get lower trading costs in government bond markets than smaller trades, consistent with the previous literature on the "size discount" studied in corporate bond and municipal bond markets. Second, our nonanonymous data set allows us to decompose the cost-size relation into within-client and cross-client variation. We find that trading costs increase in trade size once we control for client identity, generating a "size penalty."

These two findings are illustrated in Figure 1, which shows the relationship between trade size and trading costs in government bonds from two different model specifications. The left panel of Figure 1 plots the fitted linear regression line from a pooled regression of trading costs

[1] See Green, Hollifield, and Schurhoff (2007b) and the related literature.

[2] Easley and O'Hara (1987), Glosten and Milgrom (1985), and Kyle (1985), among others.

[3] See Ho and Stoll (1981) and Biais (1993) and the related literature.

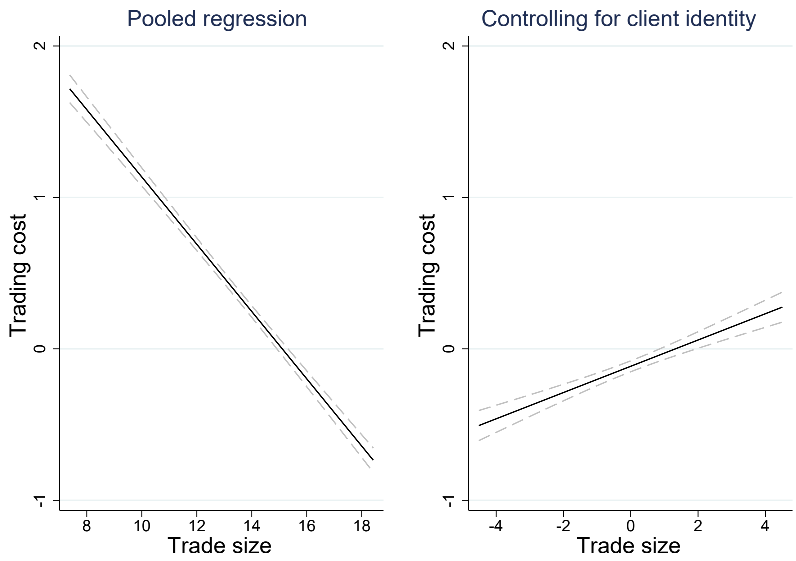


Figure 1
The relation between trade size and trading costs: The role of traders' identity
The figure illustrates the relation between trade size and trading costs on the U.K. government bond market, covering the period 2011m8-2017m12. The figure shows a linear regression line on the pooled, transaction-level data (left panel) and on the data after we removed client-specific averages from trading costs and trade sizes corresponding to each trade (right panel). Trading costs are measured by equation (1) (building on O'Hara and Zhou (2021)), and trade size is measured by the natural logarithm of the trade's notional. The estimated regression lines are based on around 1.2 million observations. The confidence bands are based on 95% standard errors as in Gallup (2019).

on trade size. The trade-level regression shows that larger trades incur lower trading costs, consistent with the findings of size discount in other OTC markets.[4] Our novel contribution is to isolate the within-client variation in the cost-size relation: the right panel of Figure 1 shows the regression line after removing the client-specific average from trading costs and trade size, showing evidence on size penalty. This suggests that the size discount is driven by the cross-client variation of larger traders facing lower trading costs and trading larger amounts, whereas the size penalty is driven by the within-client variation of the *same trader* facing higher trading costs on larger trades. This will be shown rigorously by regression analysis further below.

[4] For evidence on the size discount in the U.S. corporate bond market, see Schultz (2001), Bessembinder, Maxwell, and Venkataraman (2006), Edwards, Harris, and Piwowar (2007), Goldstein, Hotchkiss, and Sirri (2007), Hendershott and Madhavan (2015), and O'Hara, Wang, and Zhou (2018), among others. Similar evidence from the U.S. municipal bond market is presented by Harris and Piwowar (2006) and Green, Hollifield, and Schurhoff (2007a, 2007b), among others.

Third, we exploit cross-sectional variation in client types and find that the size penalty is larger for more sophisticated clients (hedge funds and asset managers), and it is smaller for less sophisticated clients (pension funds, foreign central banks, insurance companies, etc.). Fourth, we additionally exploit time-series variation in the magnitude of macroeconomic surprises and find that the size penalty, faced by more sophisticated clients, is larger during informationally intensive periods, such as trading days, that coincide with the arrival of large macroeconomic shocks. In contrast, the size penalty faced by less sophisticated clients is similar across trading days regardless of the magnitude of macroeconomic shocks at the time. Fifth, we also exploit cross-market variation by identifying clients who simultaneously trade in government bonds as well as in corporate bonds. We find that the size penalty is larger in corporate bonds than in government bonds, and, importantly, this difference is more pronounced amongst more sophisticated clients.

Taken together, we interpret these results as evidence that information-based explanations contribute to the heterogeneity in size penalty. To the extent that more sophisticated clients are more likely to trade on information than less sophisticated clients, the differential degree of size penalty across client types, implied by the difference-in-differences approach, is consistent with theories of asymmetric information (Glosten and Milgrom 1985). The triple differences approach, that uses time variation in the magnitude of macroeconomic surprises, corroborates this interpretation.

The triple differences approach using cross-market variation shows that other, inventory- and liquidity-based factors are also likely to play a role, insofar as the size penalty is larger in corporate bonds than in government bonds regardless of client types. While corporate bonds are informationally more sensitive assets than government bonds (Brancati and Macchiavelli 2019; Arnold and Rhodes 2021), liquidity and interdealer intermediation is also considerably smaller in the U.K. corporate bond market than in the government bond market. This makes it more costly for corporate bond dealers to execute trades with large size (Chen, Lesmond, and Wei 2007), which could explain the larger size penalty in corporate bonds. However, assuming that the liquidity effect of a large corporate bond trade should be the same *regardless of client type* (e.g., the sophistication of the client initiating the trade), the larger increase in size penalty among more sophisticated clients is consistent with the presence of informational channels over and above what is explained by inventory-based mechanisms.[5]

Our sixth result is that the trading activity of more sophisticated clients is a strong predictor of future returns when these clients trade in larger sizes. In contrast, we do not observe such a positive correlation between trade size and future returns in the case of less sophisticated clients. This reinforces that the size penalty captures a significant information component. To study the nature of information, proxied for by trade size, we test whether information pertains to future order flow or to learning about value-relevant information. such as the processing of macroeconomic news (Farboodi and Veldkamp 2020). We find evidence for both channels.

In our empirical design, we use various combinations of fixed effects to control for other forces that may drive the cost-size relation (though we shall acknowledge the limitations of the fixed effect approach given the complexity of unobservable factors at play). For example, the size penalty also can be driven by an inventory imbalance channel: a large-sized trade is more likely to cause skewed dealer inventory imbalance. Therefore, the dealer would be forced to cover its resultant inventory cost by charging a higher trading cost (Ho and Stoll 1981), generating a size penalty. While this force is likely to be present in the data, there are at least two reasons we interpret our results as being driven by additional, information-based factors over and above this inventory channel. First, our trade-level regressions include dealer-day fixed effects that control for the linear effects of any daily shock to dealers' inventory that could drive the cost-size relation. Second, inventory-based mechanisms alone are less likely to explain the heterogeneity in the size penalty across more sophisticated and less sophisticated clients. However, if dealers' inventory cost functions are sufficiently convex and if more sophisticated clients systematically trade larger amounts than less sophisticated clients, then the heterogeneity in the size penalty across client types could be explained by inventory channels. However, our sample does not support this possibility; instead we find that, if anything, more sophisticated clients seem to trade in smaller sizes.

Moreover, the size penalty could be affected by the strength of the trading relationship between clients and dealers (Di Maggio, Kermani, and Song 2017; Hendershott et al. 2020). While in our baseline regression, we use client-dealer fixed effects to control for the average effect of relationships on trade size and trading costs, we subsequently explore the role of relationships in shaping the size penalty. At least

[5] In a similar spirit, we also revisit the cost-size relation during the COVID-19 period (Appendix Section E), which additionally provides an ideal setting for performing a cross-check using another, more recent data sample (2018-2020).

two mechanisms that could generate a lower size penalty between counterparties that have stronger relationships. First, informational asymmetry is likely to be lower when the dealer and the client has a stronger relationship. Second, dealers may give preferable prices to certain clients with the expectation to learn from them.[6] If this incentive to learn is stronger during more turbulent periods (when the given clients also trade in larger size), then the size penalty could be partially offset. Consistent with this, our evidence shows that the size penalty is stronger on trades that are between counterparties with a weaker trading relationship.

Our empirical results highlight that controlling for traders' identity is crucial for understanding trading costs in nonanonymous OTC markets. In centralized exchanges, where client identity is not revealed before the trade, client identity is not relevant for the estimation of trading costs. However, in OTC markets, client identity is observable to dealers and naturally enters dealers' pricing function. Without information on clients' identity and by simply looking at the relationship between trade size and trading costs in a pooled regression, one would likely underestimate the marginal cost faced by market participants when they increase the size of their trades.

We make two contributions to the literature discussed below. First, we provide trade-level evidence on the size penalty in multiple OTC markets, thereby adding to the existing literature which typically focused on aggregate order-flow analysis to study the link between bond prices and orderflow.[7] Second, we exploit a number of unique sources of variation in our nonanonymous data set to isolate the role of information asymmetry from other factors in driving the size penalty. This reconciles a tension between the theoretical prediction of size penalty in the asymmetric information literature and the empirical pattern of size discount consistently documented by previous studies.

Our paper builds on two mains strands of the empirical literature. First, we draw on previous studies on the determinants of trading costs in corporate (Schultz 2001; Bessembinder, Maxwell, and Venkataraman 2006; Edwards, Harris, and Piwowar 2007; Goldstein, Hotchkiss, and Sirri 2007; Feldhutter 2012; Hendershott and Madhavan 2015; O'Hara and Zhou 2021) and municipal bond markets (Harris and Piwowar 2006; Green, Hollifield, and Schurhoff 2007a, 2007b; Li and Schürhoff

[6] See Leach and Madhavan (1993), Ramadorai (2008), Osler, Mende, and Menkhoff (2011), Golosov, Lorenzoni, and Tsyvinski (2014) and Pinter, Wang, and Zou (2020) and the related literature on the link between dealers' price setting behavior and learning.

[7] See Brandt and Kavajecz (2004), Pasquariello and Vega (2007), and Fleming, Mizrach, and Nguyen (2018), among many others.

2019).[8] We contribute to this literature by isolating the role of client identity in driving the relationship between trading costs and trade size, and to combine this client-level heterogeneity with other variations in our unique data set to develop our empirical tests. Second, we contribute to the empirical literature on informed trading in government and corporate bond markets (Brandt and Kavajecz 2004; Green 2004; Pasquariello and Vega 2007; Fleming, Mizrach, and Nguyen 2018; Hendershott, Kozhan, and Raman 2020; Czech et al. 2021; Kondor and Pinter 2022). Compared to these studies, we focus on how analyzing the cost-size relation in bond markets can reveal the presence of informed trading.

Our empirical results are able to inform the theoretical literature on OTC markets. Specifically, our evidence on the size penalty is consistent with previous models featuring asymmetric information (Glosten and Milgrom 1985; Kyle 1985; Easley and O'Hara 1987). Our evidence on the size discount is consistent with clients facing price discrimination from dealers, possibly because of the heterogeneity in their bargaining power or search intensity (Duffie, Gârleanu, and Pedersen 2005; Green, Hollifield, and Schurhoff 2007b; Pinter and Uslu 2021).

1. Measurement and Main Hypotheses

1.1 Data

To distinguish between the roles of trade size and trader size in bond markets, one needs a detailed transaction-level data set which contains information on the identity of both sides of a trade. The ZEN database sourced by the U.K. Financial Conduct Authority, contains this information along with information on the transaction time, the transaction price and quantity, the International Securities Identification Number, the account number, and buyer-seller flags. Our sample covers the period between August 2011 and December 2017. Our analysis focuses on transactions that occur between clients and designated market makers, called Gilt-Edged Market Makers (GEMMs). GEMMs are the primary dealers in the U.K. government bond market, and the majority of client-dealer trades are intermediated by them.[9] After filtering out all duplicates, erroneous entries, we are left with approximately 1.2 million observations for government bond market trades and about the same number of observations for corporate bond

[8] See Hotchkiss and Jostova (2017), Biais and Green (2019), and Bessembinder, Spatt, and Venkataraman (2020) for recent surveys.

[9] GEMMs acts as the counterparty for about 65% of all client trades in our sample, but their intermediation activity is lower in corporate bonds. For further details on the identities that make up GEMMs, see https://www.dmo.gov.uk/responsibilities/gilt-market/market-participants/.

market trades. Further details on the data construction can be found in Internet Appendix B.

A key aspect of our empirical analysis is that we are able to see the identity of both counterparties for each transaction, a unique feature of the ZEN database also used in Czech et al. (2021) and Kondor and Pinter (2022). Following these papers, we distinguish between more sophisticated clients (hedge funds and asset managers) and less sophisticated clients (pension funds, foreign central banks, commercial banks, international policy institutions, insurance companies, nonfinancial investors). This classification is motivated by the recent evidence on the enhanced ability of more sophisticated clients to predict future bond returns (Czech et al. 2021). We identify around 600 clients that cover more than 90% of the trading volume between clients and dealers.

1.2 Institutional details

Our sample contains trade reports on both conventional gilts and inflation-linked gilts that are issued by Her Majesty's Treasury (HMT) on behalf of the U.K. government. While these bond are listed on the London Stock Exchange (LSE), the majority of secondary market trading occurs off-exchange, through bilateral transactions among client and dealers, facilitated by phone calls or electronic trading platforms. In our sample, the fraction of volume in the client-dealer segment that occurs on the exchange and over-the-counter is about 15% and 85%, respectively.[10]

At the core of the gilt market's functioning are the Gilt-Edged Market Makers (GEMMs), designated as primary dealers by the U.K. Debt Management Office (DMO) responsible for U.K. government debt management. Their number hovers around 18 during our sample, and there are a few instances of entry and exit during this period.[11] GEMMs are crucial for providing continuous, responsive two-way prices, ensuring liquidity in the secondary gilt market. Operating during regular business hours, they actively make markets and quote prices for customers. The spread between their bid and ask prices is expected

[10] Around 82% of the over-the-counter trading volume occur via bilateral transactions, with the remaining 18% going through trading platforms. We do not see a pronounced difference between more or less sophisticated clients in terms of their venue choice, though the share of volume on the exchange as well as via bilateral transactions is somewhat larger for the group of more sophisticated clients, whereas less sophisticated clients trade somewhat more on the platform.

[11] For example, see the exit of Credit Suisse and Societe Generale as primary dealers and the entry of Lloyds.

to be reasonable, though not rigidly defined due to market conditions' impact on spreads.[12]

Enjoying privileges for their market-making role, GEMMs possess exclusive participation rights in DMO-run gilt primary auctions. They also receive a noncompetitive allowance of 15% of auctioned debt (DMO 2021). In the secondary market, GEMMs hold a preferred counterparty status: in almost all of its gilt market operations, the DMO transacts only with the GEMMs. Regulated by the U.K. Financial Conduct Authority (FCA), GEMMs are obliged to report all their secondary-market gilt trades to the FCA.[13] In addition, GEMMs are strongly encouraged to provide comprehensive and accurate real-time price information to their client bases, through either their dealer-to-customer platforms or multidealer electronic trading platforms or exchanges (DMO 2021).

While GEMMs play a quantitatively important role in intermediating corporate bonds as well, they do not enjoy the same privileges for their market-making role as in the gilt market, and they also face large competition from other market-makers in this market. As reported by Pinter and Uslu (2021), over 90% of the trades (in terms of trading volume) are intermediated by GEMMs in the client-dealer segment of the government bond market, while this share amounts to a still sizeable 85% in the corporate bond market. In terms of number of transactions, GEMMs intermediate about 89% and 63% of trades in the government and corporate bond markets, respectively.[14] While corporate bonds also trade on electronic platforms and on a central exchange, the relative share of these trading venues (compared to bilateral trading) is smaller than in the case of gilt trading.

1.3 Measurement and summary statistics

One of the two key variables in our empirical analysis is trade size, which we measure as the pound value of the given trade's notional. We will use the natural logarithm of trade size as the key independent variable in our regressions below; hence we briefly describe its empirical

---

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

[13] In addition, central to the gilt market structure, Interdealer Brokers (IDBs) act as intermediaries exclusively between GEMMs, allowing anonymous transactions. This setup preserves confidentiality and safeguards inventory management, which is crucial for market liquidity. Operating on a matched principle basis without proprietary positions, IDBs facilitate smooth transactions. Alongside IDBs, Agency Brokers aid trades between dealers and end-investors. See DMO (2011) for further details.

[14] The lower share of GEMMs in corporate bond intermediation underscores that GEMMs face fiercer market-making competition from non-GEMMs (e.g., other large clients and dealer banks) in this market compared to the gilt market. This is consistent with the structural estimation result of Pinter and Uslu (2021) that GEMMs have lower market power in the corporate bond market compared to the gilt market.

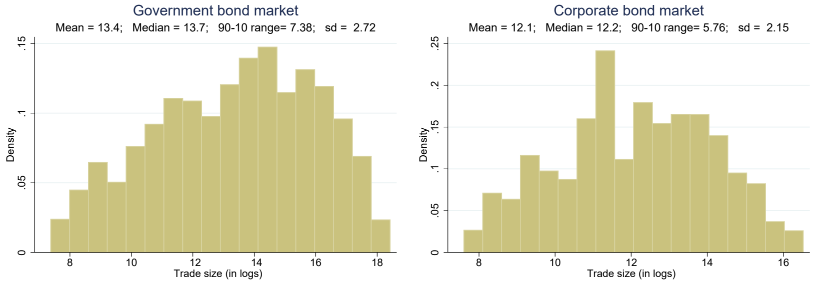


Figure 2
Trade size distributions
These figures summarize the size distributions in the U.K. government bond (left panel) and corporate bond (right panel) markets, based on trade-level data spanning the period 2011m8-2017m12. The construction of the histograms is based on the trade-level data set after trimming it at the 1%--99% level.

distribution here. Figure 2 illustrates the (log) size distribution for the U.K. government bond and corporate bond markets, along with selected summary statistics. Both the mean and the median values are larger in government bonds (13.4 and 13.7) than in corporate bonds (12.1 and 12.2). The dispersion of the distribution, measured in standard deviation, is about 0.5 log points larger in government bonds than in corporate bonds (2.72 vs. 2.15).

Table G.1 in the Internet Appendix provides further summary statistics, showing that the mean and median trade size for government bonds in our sample is about £6.3 million and £0.86 million, which suggests a sizeable skew in the distribution. While trades in corporate bonds tend to be considerably smaller, we do not see a discernible difference in the size distribution across more and less sophisticated clients.

The second key variable in our empirical analysis is trading costs, which we measure by following O'Hara and Zhou (2021). Specifically, for each trade v we compute the following measure:

$$Cost_v = \left[\ln(P_v^\star) - \ln\left(\overline{P}\right)\right] \times \mathbf{1}_{B,S}, \tag{1}$$

where $P_v^\star$ is the transaction price, $\mathbf{1}_{B,S}$ is an indicator function equal to one when the transaction is a buy trade, and equal to -1 when it is a sell trade, and $\overline{P}$ is a benchmark price, which in our baseline is the average price of all transactions in bond k on trading day t. We multiply $Cost_v$ by 10,000 to compute costs in basis points of value. As shown below, our baseline results are robust to using four alternative ways to compute $\overline{P}$.[15] Note that transaction prices in the ZEN data set do not include

[15] The four alternatives are as follows. First, we compute $\overline{P}$ as the average transaction price in bond k, trading day t, and dealer j. Second, we compute $\overline{P}$ as the average transaction

commission (or any accrued interest when relevant); hence our measure of trading cost (1) can be seen as a lower bound on the total cost of trade execution.[16]

Table 1 provides summary statistics of our baseline measure of trading costs using our sample of government bond trades. We summarize the medians of our cost measure after double sorting the data along different trade size and maturity categories. Specifically, we follow O'Hara, Wang, and Zhou (2018) in creating four size categories: micro (£1-£100,000), odd-lot (£100,000-£1,000,000), round-lot (£1,000,000-£5,000,000) and block (above £5,000,000). We also sort trades into three maturity categories: short and medium (0–8 years), long (8–20 years), and very long (above 20 years) maturities.

Importantly, we compute the median cost measures for the different size-maturity category pairs using two different specifications. Panel A of Table 1 reports medians using the pooled data set, whereas panel B reports medians after we remove client-specific averages from trading costs, that is, after purging out client fixed effects from our cost measure. The two different specifications are meant to separate the role of *trader size* from the role of *trade size* in driving the cost-size relation, thereby highlighting the role of client identity in affecting trading costs.

Panel A shows that trading costs are largest for micro trades, and costs monotonically fall in trade size, with the exception of block trades that incur a higher cost than round-lot trades across all maturities. By far, the highest trading costs in all maturity categories concentrate in micro trades. With the exception of block trades, these results are consistent with the size discount which has been documented, using similarly pooled data sets, in other fixed income markets (e.g., Edwards, Harris, and Piwowar 2007; Harris and Piwowar 2006).

Panel B of Table 1 reports the medians after controlling for client identity (with fixed effects) in trading cost measurement. Results from this specification imply that a client tends to face lower trading costs in smaller (micro and odd-lot) trades than larger (round-lot and block) trades across all maturity categories. Overall, panel B highlights that the within-client variation in trading costs seems to flip the sign of the cost-size relation, thereby giving rise to a size penalty. While these summary statistics add to the evidence illustrated by Figure 1, a number of other factors could be affecting both trade size and trading costs, including

price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.-3 p.m., or (3) after 3 p.m. Third, we also compute $\overline{P}$ as the average transaction price in bond k, trading day t, separately for buy and sell trades. Fourth, we also compute $\overline{P}$ as the average transaction price in bond k, trading day t, using only trades on the interdealer market.

[16] For further details, see Section 7.15. of the Transaction Reporting User Pack for the ZEN data set: https://www.fca.org.uk/publication/finalised-guidance/fg15-03.pdf.

Table 1
Summary statistics: Trading costs, trade size, and maturity

(a) Results from the pooled data set

	Trading cost			Number of obs.		
Maturity	0-8Y	8-20Y	>20Y	0-8Y	8-20Y	>20Y
Trade-size category						
£1–£100,000	0.778	1.447	0.530	157,533	87,631	89,004
£100,000–£1,000,000	-0.117	-0.097	-0.289	123,848	89,900	107,341
£1,000,000–£5,000,000	-0.267	-0.374	-0.420	102,478	91,682	81,865
>£5,000,000	-0.021	-0.171	0.200	145,555	108,783	88,675

(b) After removing client-specific averages

	Trading cost			Number of obs.		
Maturity	0-8Y	8-20Y	>20Y	0-8Y	8-20Y	>20Y
Trade-size category						
£1–£100,000	-0.627	-0.115	-0.072	157,533	87,631	89,004
£100,000–£1,000,000	-0.528	-0.393	-0.395	123,848	89,900	107,341
£1,000,000–£5,000,000	-0.240	-0.051	-0.173	102,478	91,682	81,865
>£5,000,000	0.146	0.048	0.381	145,555	108,783	88,675

The panels present medians of our baseline measure of trading costs (1), measured in basis points, for different segments of the maturity and size distributions. To calculate transaction costs, we use the benchmark price computed as the average price of all transactions at the bond-day level. Panel A shows median trading costs using the pooled data set of government bond trades. Panel B shows the median trading costs after removing the client-specific mean (i.e., purging out client fixed effects) from trading costs. The sample covers the period 2011m8–2017m12.

time variation in market conditions, client-dealer relationships among others. We therefore turn to regression analysis in the next section to control for these factors and to uncover the possible mechanisms that drive the size penalty.

1.4 Main hypotheses

In this section, we motivate and summarize our main testable hypotheses. As referenced in our Introduction, there is ample evidence from various OTC markets on the size discount, which is often rationalized with theories of search and bargaining frictions. These models (e.g., Green, Hollifield, and Schurhoff 2007a) give rise to a negative cross-sectional relationship between trade size and trading costs, as larger traders (who tend to trade larger amounts) have higher bargaining power and can thereby achieve lower transaction costs compared to smaller traders.

In contrast, other forces, such as asymmetric information and inventory imbalances, could generate a size penalty in OTC markets. For instance, a size penalty could arise from information asymmetry if a sophisticated hedge fund trades larger amounts after receiving an informative signal (e.g., Kyle 1985), or from inventory concerns if it is more costly for a dealer to take a larger order from its client (e.g., Ho and Stoll 1981). One could therefore expect that controlling for client identity (thereby controlling for the size discount) would lead to a positive cost-size relationship in the data.[17]

Moreover, more sophisticated clients (e.g., hedge funds) are more likely to trade on private information, subjecting dealers to more information asymmetry, while less sophisticated clients (e.g., pension funds) are more likely to trade for hedging purposes.[18] On the other hand, a dealer facing clients who seek to offload a large quantity of assets would be subject to the same inventory cost, irrespective of whether the client is more or less sophisticated. Cross-sectional variation in client types would therefore generate heterogeneity in the size penalty that is less likely to be explained by dealers' inventory imbalance. Given that asymmetric information is increased during informationally intensive periods, such as the arrival of public news,[19] it is also reasonable to expect an increase in the size penalty during these periods, which is larger among more sophisticated clients. We summarize these predictions in the following hypotheses:

Hypothesis 1 (Size discount). Trading costs are smaller for larger trades, and this is driven by the cross-client variation of larger clients facing lower costs than smaller clients.

Hypothesis 2 (Size penalty). A given client faces higher trading costs on larger trades than on smaller trades.

Hypothesis 3 (Client heterogeneity). The size penalty is bigger for more sophisticated clients compared to less sophisticated clients.

Hypothesis 4 (Macro news). The size penalty is larger during big macroeconomic news, and this difference is more pronounced among more sophisticated clients.

---

[17] Note that Section A in the Internet Appendix presents a simple model that includes both bilateral bargaining and information asymmetry to formalize the coexistence of the (cross-client) size discount and the (within-client) size penalty.

[18] See Czech et al. (2021) and Kondor and Pinter (2022) for recent evidence.

[19] See Green (2004) and Pasquariello and Vega (2007), among others.

We now introduce our empirical design before presenting evidence that supports our main hypotheses.

2. Empirical Analysis

2.1 The empirical model

Our baseline specification is the following trade-level regression:

$$Cost_v = \beta \times Size_v + \alpha_{k,t} + \lambda_{i,m} + \mu_{j,t} + \delta_{i,j} + \varepsilon_v, \quad (2)$$

where $Cost_v$ is the trading cost as computed in 1, $Size_v$ is the natural logarithm of the given trade's notional (in £s) and our control set includes combinations of fixed effects at the levels of client i, dealer j, bond k, day t and month m. The key object of interest is the estimated value of β: if trading cost is the same for large and small trades, then we would not expect β to be significantly different from zero.[20]

Regarding our regression design, a brief discussion on endogeneity is warranted as our approach (2) is to effectively regress prices on quantities. We acknowledge that we do not aim to estimate the *causal* effect of trade size on transaction costs and the main drivers of trade size and trading costs are unobservable primitives (such as private information). Our objective is to isolate the contribution of information-based mechanisms (Kyle 1985) to the observed size-penalty and use additional variation in the data as well as a combination of fixed effects to disentangle information from liquidity effects and other potential confounds. Here, we discuss seven alternative mechanisms that could give rise to the size penalty, and explain how we aim to control for these mechanisms in our regressions.[21]

First, client size may change over time which could be observable to dealers. As a client's size increases, its average trade size may increase while its transaction costs may fall as its bargaining power increases and its relationships with counterparties develop. This could bias the estimated cost-size relationship downward.

Second, clients' average informativeness may change over time which might be inferred by dealers. This could increase average trade size, whereas the effect on transaction costs may be positive if the classic

[20] There are at least two reasons our baseline regression is at the trade level instead of more aggregated levels. First, our trade-level regression allows us to control for possible confounds related to bond-, dealer-level heterogeneity with multidimensional fixed effects. Second, the analysis of the cost-size relation would be more difficult at more aggregated levels, given the tendency of institutional investors to pursue order-splitting strategies (Campbell, Ramadorai, and Schwartz 2009; Kondor and Pinter 2022).

[21] We acknowledge, however, that this is not an exhaustive list, and fixed effects models like ours have constraints on what they can control for (given the interplay of unobservable factors).

adverse selection channel dominates or costs could also fall if dealers aim to attract the given client with price discounts to learn from them.[22] Hence the direction of the bias is more ambiguous in the case.

We use client-month fixed effects ($\lambda_{i,m}$) to control for these effects. The underlying assumption is that changes in both client size and average informativeness occur at lower frequency than a month. By controlling for these lower frequency mechanisms with fixed effects, we try to isolate the effects due to higher-frequency objects, such as information *shocks* or liquidity *shocks* that are unobservable to dealers, as also formalized in our theoretical model in the Internet Appendix A.

Third, time variation in dealers' balance sheet constraints (Bessembinder et al. 2018) could affect the cost-size relationship. Relatedly, variation in dealers' inventory (of all bonds) could correlate with clients' trading needs. For example, during illiquid market conditions, such as the recent LDI crisis in the United Kingdom (Pinter 2023), a dealer may face large selling pressure from clients, and tightened inventory constraints force the dealer to increase spreads. This could bias the estimated cost-size relationship upward, which we aim to control for with dealer-day fixed effects ($\mu_{j,t}$).[23]

A fourth, related possibility is that variation in market conditions for a particular bond could be correlated with clients' trading needs. This could be particularly relevant after primary issuances or following (correlated) shocks to preferred habitat investors, such as pension funds and LDI investors who regularly use specific bonds to duration hedge contractual pension liabilities (Klingler and Sundaresan 2019). This may bias the estimated cost-size relationship upward, which we aim to control for with the inclusion of bond-day fixed effects ($\alpha_{k,t}$).

Fifth, the strength of the trading relationship between client and dealer could affect both the average trade size and transaction cost (Di Maggio, Kermani, and Song 2017). For example, a stronger relationship may lead to larger trade size and lower transaction costs paid by the client, which could bias the estimated cost-size relationship downward. We aim to control for the linear effects of this mechanism with client-dealer fixed effects ($\delta_{i,j}$), which effectively allows the comparison of trades that are executed by the same counterparties in different points in time. Sixth, it is possible that a given client can be hit by liquidity shocks (e.g., because of an unexpected urgency to trade or shocks to client inventory), which could drive up both the

[22] See Glosten and Milgrom (1985), among many others, for the first mechanism and Leach and Madhavan (1993), Osler, Mende, and Menkhoff (2011), Pinter, Wang, and Zou (2020), and Ramadorai (2008) for the second mechanism.

[23] Note that dealer-day fixed effects automatically control for time-invariant dealer-heterogeneity, which has been found important in determining transaction costs (Hollifield, Neklyudov, and Spatt 2017).

client's trade size and trading cost, inducing a size penalty.[24] However, not observing a client's liquidity shock would not bias our estimate on the difference between the size penalties across client types as long as liquidity shocks have the same size and are evenly distributed across more or less sophisticated clients. This motivates the use of difference-and-differences style estimation (in Section 2.4.1) whereby we add an indicator variable (capturing client type) to our baseline model 2 and interact it with trade size.

Seventh, the arrival rate or the size of liquidity shocks might be correlated with client type, which could bias our difference-and-differences-style estimate. For example, more sophisticated clients might be better at managing liquidity shocks (e.g., better at concealing their urgency to trade or managing inventory) than less sophisticated clients.[25] To address this, we introduce the second layer of differences by additionally exploiting time-series variation in the amount of macroeconomic news (Section 2.4.2).[26] The underlying assumption is that the arrival rate or size of liquidity shocks (while potentially correlated with client types) would be uncorrelated across high and low macro news periods for the same client. If this assumption is satisfied, then the estimates given by our triple differences approach would not be biased by these liquidity-related confounds.

While our strategy to employ fixed effects can plausibly control for several confounds, we acknowledge the limitations of our approach given the complexity of unobservable factors at play.[27] Against this backdrop, we next turn to presenting our main empirical results.

2.2 The role of trader identity in the cost-size relation

Table 2 presents the results for our baseline regression 2. All regressions include bond-day fixed effects that aim to control for the linear effects of aggregate shocks that may affect bonds heterogeneously. We move gradually from the least restrictive specification (column 1) to the most restrictive specification (column 5). Consistent with the Figure 1 above, the inclusion of client fixed effects (column 2) makes the biggest change to the estimation results by flipping the sign of the estimated effect: without client fixed effects, a one log unit increase in trade size is

[24] This could be the result of a pure liquidity shock as in Vayanos and Vila (2021) or shocks to preferences as in Lagos and Rocheteau (2009).

[25] Note that in this case the omitted variable problem would, if anything, bias our first diff-and-diff estimate downward.

[26] In this spirit, we also exploit cross-market variation by comparing the cost-trade relation in the government and corporate bond markets (Section 3.5).

[27] Future research could improve on our empirical design by trying to find exogenous variation in private information or inventory constraints of clients.

Table 2
Trading costs and trade size in government bond markets: The role of traders' identity

	(1)	(2)	(3)	(4)	(5)
Trade size	-0.217***	0.102***	0.121***	0.134***	0.158***
	(-4.04)	(3.05)	(4.01)	(4.25)	(5.19)
N	1,274,295	1,274,289	1,274,289	1,269,855	1,269,238
R^2	.055	.061	.062	.134	.139
Day#Bond FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer FE	No	No	Yes	No	No
Day#Dealer FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects (2). The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. *t*-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels.* $p < .1$; ** $p < .05$; *** $p < .01$.

associated with a size discount of -0.217 bps. In contrast, the inclusion of client fixed effects results in a size penalty of 0.1 bps.

In column 3, we include dealer fixed effects to control for (time-invariant) dealer-heterogeneity in determining transaction costs. In column 4, we include dealer-day fixed effects to control for time variation in the tightness of balance sheet constraints of dealers and client-month fixed effects to control for lower frequency variation in client characteristics. In column 5, we include client-dealer fixed effects to control for the linear effects of client-dealer relationships. We find that the results are qualitatively robust to including these additional fixed effects. Note that results are also similar when estimating an alternative version of model 2, whereby we explicitly control for observable characteristics instead of using fixed effects (Internet Appendix C).[28]

In terms of economic significance, the estimated effects are nonnegligible. The estimated coefficient in column 5 of Table 2 together with the fact that the standard deviation of trade size is around 2.72 log point (Figure 2) suggest that a one-standard-deviation increase in trade size increases trade costs by about 0.43 bps. This is sizeable compared

[28] Table C.1 gives qualitatively similar results to our baseline Table 2, though yielding lower explanatory power in terms of R^2 statistics. The average of a client's monthly trade volume achieves a similar result to using a client fixed effect insofar as the estimated cost-size relationship turns from negative (column 1 of C.1) to positive (column 2 of C.1). That one can replace the client fixed effect with the client trade-volume to obtain the size penalty is consistent with Hypothesis 1, which says that the size discount is driven by the cross-client variation of larger clients facing lower costs than smaller clients.

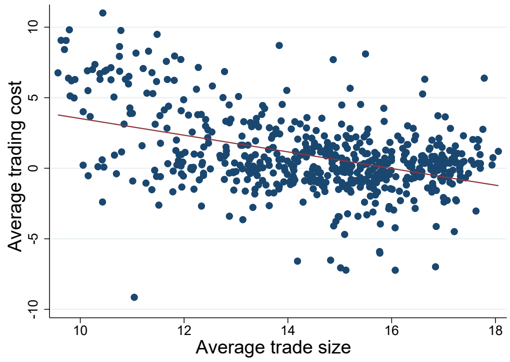


Figure 3
Trading costs and trader size in the cross-section
This figure shows a scatter plot of average client trading costs (vertical axis) against average trade size (horizontal axis) at the client level in the U.K. government bond market. Average trading cost is the client-specific mean of our baseline cost measure (1). Average trade size is the natural logarithm of the average nominal size of a client's transactions. To reduce noise, we trim the data set at 1% level, leaving 586 observations. The estimated slope is $\hat{\gamma} = -0.59$ with a t-statistic (based on robust standard errors) of -9.6.

to average bid-ask spreads of around 0.25 bps and 0.33 bps in U.K. gilts of 10- and 30-year maturities, respectively, during this period.[29]

2.3 The size discount in the cross-section

We argued above that the size discount implied by pooled regressions (left panel of Figure 1 and panel A of Table 1) is driven by the cross-client variation of larger traders facing more favourable trading costs and trading larger amounts. To show this rigorously, we collapse our data set at the client level and estimate the following cross-sectional regression for client i:

$$Cost_i = c + \gamma \times Size_i + \varepsilon_i, \qquad (3)$$

where $Cost_i$ is the unweighted mean trading cost (1) based on all trades of client i, and $TraderSize_i$ is measured as the natural logarithm of mean trade size of client i.

The results are summarized in Figure 3, confirming a statistically significant size discount in the cross-section, which is consistent with Hypothesis 1. In spite of the simplicity of the cross-sectional regression (with various other dimensions of client heterogeneity not featured), the estimated model delivers a nonnegligible R^2 of .2. These results are robust to using an alternative measure of trader size, such as clients' total monthly trading volume averaged across months. Moreover, the

[29] The computation of bid-ask spreads is based on daily average quote data from Bloomberg and covers the period 2012-2017.

baseline scatter plot looks similar when we control for clients' average monthly dealer connections (Hendershott et al. 2020; Kondor and Pinter 2022) as well as average monthly intensity – measured as the natural logarithm of total number of transactions – of client i (O'Hara, Wang, and Zhou 2018).[30]

As reviewed in the Introduction, a voluminous literature documented on the size discount as an important feature of bond trading. Figure 3 adds to this literature by isolating the source of variation in the trade-level data (i.e., the cross-client variation) that drives the documented size discount.

2.4 The size penalty: Exploring the mechanisms

We now take a closer look at the within-client pattern of size-penalty documented in Section 2.2.

2.4.1 The role of trader sophistication. To explore the role of heterogeneity in client types in driving the size penalty, we first divide clients into two groups based on whether a given client is of a more sophisticated type (asset manager or hedge fund) or of a less sophisticated type (pension funds, foreign central bank, etc.). As argued, the underlying assumption is that former group is more likely to trade on information, than the latter group.[31] For these two groups $g=\{g_1,g_2\}$, we estimate an extended version of our baseline regression 2, as follows:

$$Cost_v=\sum_{w=1}^{2}\eta_w\times\mathbf{1}[i\in g_w]\times Size_v+FE+\varepsilon_v, \tag{4}$$

where $\mathbf{1}[i\in g_w]$ is an indicator function equal to one if client i belongs to group w, and zero otherwise. We present the estimates of η_1 and η_2 adjacent to each other in the regression tables and present results for tests of equality of the two coefficients.

Table 3 shows the results for the case when we estimate regression 2 for the more sophisticated and less sophisticated clients separately. This difference-in-differences (DID) approach reveals that while the cross-sectional phenomenon of size discount is present for both sets of clients (with the estimated coefficients being similar), the inclusion of client fixed effects generates a larger size penalty for the group of more sophisticated traders. The most conservative specification (column 5) shows that the size penalty is almost twice as large among more sophisticated clients (0.197) than among less sophisticated clients (0.106), consistent with Hypothesis 3.

[30] These results are shown by Figures H.1 and H.2 in the Internet Appendix.

[31] See Czech et al. (2021) for further details on the ability of more sophisticated clients to forecast bond returns.

Table 3
Trading costs and trade size: More sophisticated clients versus less sophisticated clients

	(1)	(2)	(3)	(4)	(5)
Less sophisticated clients					
Trade size	-0.217***	0.057	0.067	0.080*	0.106**
	(-4.14)	(1.22)	(1.46)	(1.79)	(2.29)
More sophisticated clients					
Trade size	-0.213***	0.140***	0.169***	0.183***	0.197***
	(-2.67)	(3.34)	(4.83)	(5.08)	(5.61)
p-values, eq. of coeff.	.966	.185	.076	.069	.115
N	1,271,112	1,271,106	1,271,106	1,264,580	1,263,963
R^2	.100	.106	.107	.202	.207
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer#ClientType FE	No	No	Yes	No	No
Day#Dealer#ClientType FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. * $p<.1$, ** $p<.05$, *** $p<.01$.

To the extent that more sophisticated traders are more likely to trade on information, heterogeneity in the degree of size penalty across the two groups of traders is suggestive of an information-based explanation (see, e.g., Kyle 1985; Easley and O'Hara 1987). However, more sophisticated clients could still face a steeper cost-size trade-off for noninformational reasons (e.g., mechanisms related to dealer inventory). Therefore, to isolate more rigorously the role of information in driving the size penalty, the next subsections extend the DID approach by adding one additional layer of "differences" related to macroeconomic surprises.

2.4.2 The size penalty around macroeconomic announcements. In this section, we estimate the role of macroeconomic announcements in affecting the degree of size-penalty. According to our Hypothesis 4, the release of large unexpected macroeconomic news leads to higher probability of informed trading (Bernile, Hu, and Tang 2016; Du, Fung, and Loveland 2018), so it increases the size penalty of both more and less sophisticated clients. Moreover, since more sophisticated clients have a higher likelihood to possess private information or a more accurate private interpretation of public signals, the increase in size penalty should be larger among this group of clients.

We build on the high-frequency methodology of Swanson and Williams (2014) to identify trading days when the surprise component of U.S. and U.K. macroeconomic announcements were unusually high.[32] Specifically, we sort trading days into two groups $s=\{s_1,s_2\}$, based on whether the magnitude of the surprise on day t was smaller or bigger than the sample median. We estimate a modified version of our baseline regression 4, as follows:

$$Cost_v=\sum_{w=1}^{2}\sum_{z=1}^{2}\eta_{w,z}\times\mathbf{1}_t[t\in s_z,i\in g_w]\times Size_v+FE+\varepsilon_v, \qquad (5)$$

where $\mathbf{1}_t[t\in s_z,i\in g_w]$ is an indicator function equal to one when a given trading day t belongs to group z and client i belongs to group w, and is equal to zero otherwise; the term FE includes various combinations of fixed effects discussed above.

The results are shown in Table 4. The size penalty faced by less sophisticated clients is virtually the same regardless of whether trading days are hit by small or large macroeconomic shocks. In contrast, the size penalty continues to be more statistically significant for sophisticated clients. Importantly, the point estimates are around 30% larger on trading days with big macroeconomic surprises (0.221) compared to days with small surprises (0.171) in the most conservative specification (column 5). We acknowledge that the statistical significance of these differences is somewhat modest, with the corresponding p-values being in the range .08-.15.

The third layer of the triple differences approach, represented by the time variation in the magnitude of macroeconomic surprises, provides a stronger evidence for the presence of information-based drivers of the size penalty, compared to the difference-in-differences approach of the previous subsection.[33,34]

An alternative interpretation of these results, however, is that the trading activity of clients during informationally sensitive periods may be endogenous to client type, for example, less sophisticated clients may refrain from trading (large amounts) compared to more sophisticated clients during high-surprise days. This selection effect combined with the possibility of dealers facing more inventory risk during high-surprise days may be driving the results in Table 4. We try to address this

[32] Our data set is obtained from the Bank of England (building on Eguren-Martin and McLaren (2015) as used in Kondor and Pinter (2022)). The method uses historical tick data to compute the change in the 3-year interest rate in a tight window (5 minutes before and 5 minutes after) around the release of both nominal and real news from both from the United Kingdom and the United States.

[33] Tables G.5 and G.6 in the Internet Appendix show that the results are similar when we experiment with four alternative measures of trading costs as left-hand-side variables in regression 5.

[34] An additional, interesting question relates to how the size penalty behaves for the more and less sophisticated clients in the days leading up to the event and afterward. This is analyzed in the Internet Appendix Section D.

Table 4
Trading costs and trade size: Around big and small macroeconomic news

	(1)	(2)	(3)	(4)	(5)
Less sophisticated clients					
Trade size#SmallNews	0.061	0.080	0.061	0.082*	0.104**
	(1.25)	(1.57)	(1.33)	(1.70)	(2.11)
Trade size#LargeNews	0.067	0.069	0.066	0.075	0.107**
	(1.44)	(1.40)	(1.41)	(1.58)	(2.13)
p-values, eq. of coeff.	.830	.673	.871	.833	.943
More sophisticated clients					
Trade size#SmallNews	0.150***	0.161***	0.141***	0.154***	0.171***
	(4.01)	(4.03)	(3.49)	(3.62)	(4.11)
Trade size#LargeNews	0.196***	0.207***	0.202***	0.205***	0.221***
	(4.55)	(4.75)	(4.77)	(4.79)	(5.17)
p-values, eq. of coeff.	.132	.119	.086	.144	.147
N	1,182,307	1,178,836	1,179,827	1,176,302	1,175,687
R^2	.106	.136	.173	.199	.204
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes	Yes
Dealer#ClientType FE	Yes	No	No	No	No
Client FE	Yes	Yes	Yes	No	No
Day#Dealer#ClientType FE	No	No	Yes	Yes	Yes
Month#Client FE	No	Yes	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs (measured in bp) on trade size (measured as the logarithm of the nominal size of the trade in £s) interacted with indicator variables denoting whether the trading day coincides with the arrival of a large or small macroeconomic surprise compared to the median, and whether the client is more or less sophisticated. The macroeconomic surprises are constructed following the high-frequency methodology of Swanson and Williams (2014). The regression also includes various fixed effects. To reduce noise, we winsorize the sample at the 1% level. *t*-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. The *p*-values correspond to the testing for the equality of coefficients, within a given client type. * $p<.1$, ** $p<.05$, *** $p<.01$.

issue two ways. First, we document how the relative presence of more and less sophisticated clients may change during low- and high-surprise days. We measure the presence of the given client group with three variables: (1) total number of daily transactions, (2) total daily trading volume and (3) total number of unique clients from the given group type. As shown in Table G.7 in the Internet Appendix, we find all three measures increase during days with large macroeconomic surprises. Importantly, we find that this increase is similar across the two client groups, indicating that the selection issue may be less of a concern. Second, Section 3.5 will exploit cross-market variation to further analyze the source of heterogeneity in the size penalty. This alternative approach is less subject to the aforementioned selection effect, as will be discussed below, because we require clients to be present and active in both markets in a given time period.

The next section turns to analyzing the relationship between trade size and trading performance to strengthen our information-based interpretation of the size-penalty, followed by an analysis of the nature of information proxied by trade size.

2.4.3 Trade size and informed trading. To explore the possible link between trade size and informed trading, we begin by examining whether certain clients' order flows are correlated with subsequent price drifts and check whether this correlation is different in periods when clients trade in unusually large amounts.

To that end, we adopt a simple portfolio approach (Di Maggio et al. 2019; Czech et al. 2021) whereby we sort bonds on each day into three terciles based on the order flow of more sophisticated clients in the "large size" and "small size" groups, that is, bonds in the upper (lower) tercile are most heavily bought (sold) by these client-groups. We then construct a long-short portfolio that goes long on the top tercile and short the bottom tercile of the assets. Additionally, the novel aspect of our portfolio strategy is to check how the cumulative daily returns of these long-short portfolios differ across the "large size" and "small size" groups and how this difference varies across more and less sophisticated clients.

Formally, we define the relative return measure for client type $k = \{LessSoph., MoreSoph.\}$, trading day t and horizon $T = \{1,2,\ldots,30\}$ as follows:

$$ExcessReturn_{k,t}^{T} = LargeSizeReturn_{k,t}^{T} - SmallSizeReturn_{k,t}^{T}, \quad (6)$$

where $LargeSizeReturn_{k,t}^{T}$ is the cumulative return from the long-short portfolio strategy whereby we buy gilts (on day t, assuming to hold till T) that are more heavily bought (on day t) by clients k that belong to the "large size" group, that is, we mimic the behavior of these clients when they trade in unusually high trade sizes on day y; similarly, $SmallSizeReturn_{k,t}^{T}$ is the cumulative return from the long-short portfolio whereby we mimic the behavior of clients when they trade in unusually small trade sizes. The measure $ExcessReturn_{k,t}^{T}$ denotes the difference in cumulative returns from these trading strategies.

Another way of describing our strategy is that we simply add an additional layer to the portfolio analysis of Czech et al. (2021) who focus on the total order flow of a particular client group. Instead, we distinguish between order flows that are initiated by clients in periods when they trade in smaller or larger average sizes. Importantly, the notion of average trade size, which we use as a reference point to classify

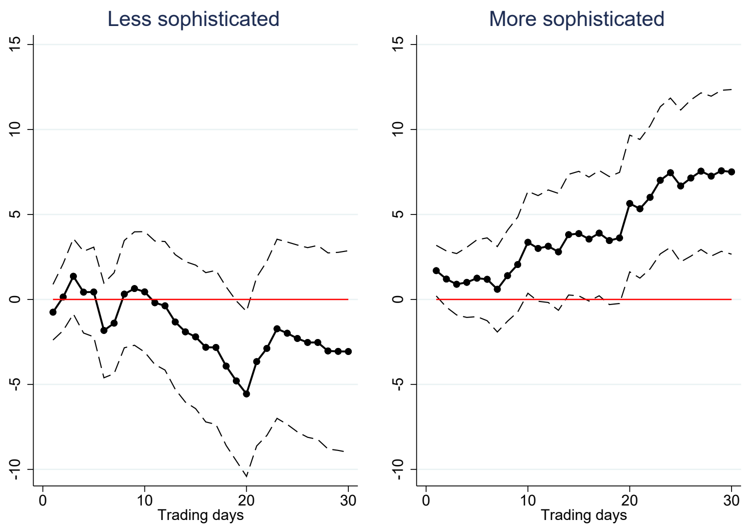


Figure 4
Long-short portfolio returns on large versus small trade size groups
This figure plots the estimated values of regression 6 over the 1-to 30-day horizon for the less sophisticated (left panel) and more sophisticated clients (right panel). The 90% confidence interval based on robust standard errors.

trades as small or large, is client-specific (which keeps this analysis consistent with our analysis of the size penalty in the earlier Section 2.2).[35]

Figure 4 shows the average values $ExcessReturn^{T}_{k,t}$ (6) for the two client groups over horizons $T = \{1,2,...,30\}$. The 90% confidence bands are based on robust standard errors that are computed based on our sample of daily return observations. We find persistently positive estimated values of $ExcessReturn^{T}_{k,t}$ in the case of more sophisticated clients without any sign of decay. The strategy yields around 8 bps at the 30-day level which is statistically significant. In contrast, we find no significant effect in the case of less sophisticated clients, and the point estimates, if anything, tend to be negative

Next, we turn to analyzing the nature of information proxied by trade size. Motivated by the recent literature (Farboodi and Veldkamp 2020; Czech et al. 2021; Kondor and Pinter 2022), we explore two hypotheses

[35] As an example, assume there are 2 clients (A and B) and 2 trading days (t_1 and t_2). Further assume that average trade size of client A is 10 and 20 on days t_1 and t_2, respectively, and the values corresponding to client B are 40 and 30 on days t_1 and t_2. According to our definition, the "small size group" would contain client A on day t_1 and client B on day t_2, and the "large size group" would contain client A on day t_2 and client B on day t_1.

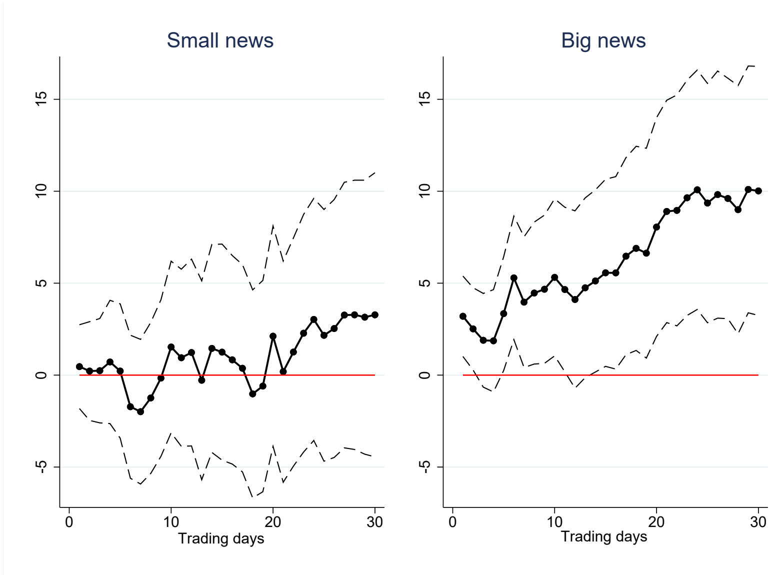


Figure 5
Decomposing the long-short portfolio returns of more sophisticated clients: Small versus large macroeconomic news
This figure plots the estimated values of regression 6 over the 1- to 30-day horizon for more sophisticated clients on days with small macroeconomic news (left panel) and on days with big news (right panel). The 90% confidence interval based on robust standard errors.

regarding the nature of information in these markets: information could pertain (1) to the prediction and processing of macroeconomic news, or (2) to the future order flow.

To provide evidence on the role of macroeconomic announcements in shaping these return patterns, we split the trading days into two equal-sized groups: one with no or small macroeconomic surprises, and another one with larger surprises. The magnitude of surprises is based on the realized high-frequency movements in asset prices around U.S. and U.K. macroeconomic announcements (Swanson and Williams 2014; Eguren-Martin and McLaren 2015) used in the previous subsection 2.4.2.

Figure 5 shows the average values of the group of more sophisticated clients on the two sets of trading days. While cumulative returns have a positive trend on both pictures, the effect is stronger and statistically significant on trading days that coincide with large macroeconomic announcements.

We now examine whether more sophisticated clients are able to better predict the future order flow of less sophisticated clients and how the potential predictability is affected by trade size. Concretely, we build on Czech et al. (2021) and estimate the following panel regression for each bond j and day t:

$$Order\ Flow^{U}_{t+1,t+d} = \alpha_t + \delta_j + \beta \times Order\ Flow^{S}_{t-4,t} + \varepsilon_{j,t}, \qquad (7)$$

where $Order\ Flow^{U}_{t+1,t+d}$ is the cumulative order flow of less sophisticated client in the next d days, $Order\ Flow^{S}_{t-4,t}$ is the cumulative order flow of more sophisticated clients in the past 4 days, α_t is a day fixed effect and δ_j is a bond fixed effect.

Importantly, we extend regression 7 by distinguishing between the order flow of more sophisticated clients depending on whether they trade in larger or in smaller sizes. Specifically, we allocate the trading days of each client into two equal-sized groups depending on whether the average trade size of the given client is larger or smaller on a trading day compared to the (client-specific) median value, that is, half the time the client would be in the “large size” group and half the time it would be in the “small size” group. We then aggregate (across the clients) the order flow of the “large size” and “small size” groups, and estimate the following regression:

$$\begin{aligned} Order\ Flow^{U}_{t+1,t+d} = {} & \alpha_t + \delta_j + \beta_1 \times Small\ Size\ Order\ Flow^{S}_{t-4,t} \\ & + \beta_2 \times Large\ Size\ Order\ Flow^{S}_{t-4,t} + \varepsilon_{j,t}, \end{aligned} \qquad (8)$$

where the coefficients of interest are β_1 and β_2 and we expect $\beta_2 > \beta_1$.

The results from estimating regressions 7 and 8 are presented in panels A and B of Table 5. Panel A of Table 5 indicates that more sophisticated clients’s orderflow in the past week significantly predicts next day’s and next week’s order flow of less sophisticated clients. These results are consistent with the findings of Czech et al. (2021). Importantly, panel B of Table 5 presents novel evidence that the majority of the predictability concentrates on those trading days when the more sophisticated clients trades in larger amounts.

3. Robustness and Extensions

3.1 Client type and trade size

A main source of variation in our empirical design is client type. A key assumption underlying the information-based interpretation of our regressions is that more sophisticated clients are different from less sophisticated clients because they are more likely to trade on information, and not because, say, they systematically trade in different quantities. To test for this, we estimate the following trade-level regression:

$$Size_v = \delta \times D_i^{Soph} + FEs + \varepsilon_v, \qquad (9)$$

where D_i^{Soph} is a dummy variable taking value one is client i is an asset manager or a hedge fund and zero otherwise.

Table 5
Predicting future order flow

	(1)	(2)	(3)	(4)	(5)	(6)
	$Order\ Flow^{U}_{t+1,t+1}$			$Order\ Flow^{U}_{t+1,t+5}$		
$Order\ Flow^{S}_{t-4,t}$	0.026***	0.026***	0.023***	0.069***	0.071***	0.055***
	(4.17)	(4.02)	(3.73)	(3.21)	(3.33)	(2.93)
N	62,249	62,249	62,249	58,761	58,761	58,761
R^2	.001	.025	.038	.001	.024	.063
Day FE	No	Yes	Yes	No	Yes	Yes
Instrument FE	No	No	Yes	No	No	Yes

(a) *The predictive power of order flow*

	(1)	(2)	(3)	(4)	(5)	(6)
	$Order\ Flow^{U}_{t+1,t+1}$			$Order\ Flow^{U}_{t+1,t+5}$		
$Small\ Size\ OF^{S}_{t-4,t}$	0.005	0.004	0.007	-0.031	-0.039	-0.025
	(0.40)	(0.34)	(0.61)	(-0.61)	(-0.76)	(-0.47)
$Large\ Size\ OF^{S}_{t-4,t}$	0.010**	0.010**	0.007*	0.050***	0.053***	0.037***
	(2.40)	(2.54)	(1.91)	(3.40)	(3.66)	(2.83)
N	57,336	57,336	57,336	54,636	54,636	54,636
R^2	.000	.027	.039	.001	.028	.064
Day FE	No	Yes	Yes	No	Yes	Yes
Instrument FE	No	No	Yes	No	No	Yes

(b) *Small-sized versus large-sized order flow*

This table presents the estimation results for regressions 7 and 8. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and bond levels. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table G.11 in Internet Appendix shows the results for the estimated values for δ using different combinations of fixed effects. The effects are statistically insignificant. The point estimates suggest that, if anything, more sophisticated clients seem to trade in smaller sizes. This rejects the possibility that heterogeneity in client types is simply picking up that more sophisticated clients trade in larger sizes, which could generate larger inventory costs for dealers and thereby larger size penalty.

3.2 Nonlinearities

Moreover, we check for nonlinearities and nonmonotonicity in the size penalty. To that end, we reestimate a variant of Table 3 by replacing *size* as a continuous variable with four dummy variables indicating which size quartile a given trade is located in, using the within-size variation of a client. As shown in Table G.12 in Internet Appendix, trading costs are the largest on the trades that are in top quartile of the size distribution, using the within-client variation of trade sizes. The results continue to be more statistically and economically significant among more sophisticated clients compared to less sophisticated clients.

3.3 Agency trades

As an additional test, we explore variation in *trade-type* to further investigate the possible information-based mechanism underlying the size penalty. In our baseline sample, about 20% of client-dealer trades are labelled as agency trades, with these trades structured as follows: trader B trades, on behalf of trader A, with trader C. In our sample, trader C is always a dealer; trader B (the agent) can be either a dealer or a client; and trader A is always a client that can be more or less sophisticated.[36] We now test whether a more sophisticated client A faces a differential size penalty when trading directly with dealers (non-agency trade) compared to trading with dealers via an agent (agency trade). The hypothesis is that if information-based mechanisms are at play, then the size penalty of the sophisticated client would be smaller on agency trades, because the given client could conceal her identity through agency.[37]

Table G.13 in the Internet Appendix shows the results for the group of more sophisticated clients from a variant of regression 4, where we interact size with a dummy variable, indicating whether or not the given trade is an agency trade. We find that the size penalty is concentrated in non-agency trades, consistent with an information-based explanation.

3.4 The role of trading relationships

While our baseline regression (2) controls for the linear effect of client-dealer relationships with fixed effects, relationship trading may affect the cost-size relationship nonlinearly. For example, to the extent that the size penalty is driven by information asymmetry, one would expect the information asymmetry to be smaller among counterparties that have a stronger relationship. In this section, we explore this possibility empirically.

To measure the role of trading relationships in affecting the cost-size relation, we first estimate the strength of client-dealer relationships as follows. Following Di Maggio et al. (2019), we compute the total trade volume for each client-dealer pair (i,m) over the previous 6 months $(t-6 \to t)$.

From the point of view of each client, we then calculate the fraction of volume effectuated with each of the dealers, denoted by the following measure:

---

[36] Our baseline sample therefore excludes the type of agency trades (typically studied in the literature) whereby (using the example above) traders A and C would be clients and trader B would be a dealer.

[37] See Smith, Turnbull, and White (2001) and the related literature.

Table 6
Distribution of trades across top client-dealer relationships

	(1) Yes	(2) No
Clients' most-favored dealer	29.55%	70.45%
Clients' two most-favored dealers	44.54%	55.46%
Dealers' most-favored client	8.60%	91.40%
Dealers' two most-favored clients	14.98%	85.02%

Column 1 of this table presents the ratio of number of transactions that are executed via favourite client-dealer relationships to the total number of transactions in our sample (with column 2 presenting the fraction of trades outside these relationships). We use four definitions of favourite relationships based on measures 10 and 11.

$$\Gamma^{c}_{i,m,t-6\to t} = \frac{\sum_{j_1}^{N_{i,m,t-6\to t}} quantity_{j_1}}{\sum_{j_2}^{N_{i,t-6\to t}} quantity_{j_2}}, \tag{10}$$

where the numerator sums across the trade quantities of client i against dealer m over the size-month period $t-6\to t$, and the denominator sums across the quantities of client i against all dealers over the same period. We then use measure $\Gamma^{c}_{i,m,t-6\to t}$ to rank dealers, with the top dealer labelled as the client's most-favored dealer.

However, a relationship that is important to a client may not be important to the dealer. Therefore, we also compute an analogous measure from the point of view of each dealer:

$$\Gamma^{g}_{i,m,t-6\to t} = \frac{\sum_{j_1}^{N_{i,m,t-6\to t}} quantity_{j_1}}{\sum_{j_2}^{N_{m,t-6\to t}} quantity_{j_2}}, \tag{11}$$

where the numerator sums across the trade quantities of dealer m against client i over the size-month period $t-6\to t$, and the denominator sums across the quantities of dealer m against all clients over the same period. We than use measure $\Gamma^{g}_{i,m,t-6\to t}$ to rank clients, with the top client labelled as the dealer's most-favored client.

Table 6 summarizes the distribution of trades in our sample across the top client-dealer relationships. We consider four cases: (1) a client's most important dealer in terms of measure 10, (2) a client's two most important dealers in terms of measure 10, (3) a dealer's most important client in terms of measure 11, (4) a dealer's two most important clients in terms of measure 11. We find that around 29.6% of trades in our sample of around 1.2 million transactions are performed between clients and their more favored dealer. The number increases to around 44.6% once we also include the client's second most important dealer in the calculation.

When we calculate these fractions from the dealers' point of view (corresponding to formula 11) the values tend to fall. This is because we have many more clients (around 600) than dealers (around 18) in our sample. As shown by Table 6, dealers perform around 8.6% of their trades with their most important client, and the fraction increases to around 15% once we include the dealers' second most important clients in the calculation.

Armed with these measures of relationship importance, we use each of the four measures to sort trades into two groups $g=\{g_1,g_2\}$, and estimate a modified version of regression 4.

Panel A of Table 7 presents the results for the case when we allow the size penalty to be heterogeneous between the clients' strongest and weakest dealer relationships. The baseline estimate (column 5 of Table 2) increases from 0.158 to above 0.2 when we consider trades between clients and their weaker dealer relationships. In contrast, the size penalty drops to around 0.07 when we consider clients' trades with their single most important dealer. This is consistent with a possibly offsetting effect of the channel proposed above (and the related literature, e.g., Leach and Madhavan 1993; Ramadorai 2008; Osler, Mende, and Menkhoff 2011; Pinter, Wang, and Zou 2020) which reduces the informational channel connected with the size penalty. Moreover, the large size penalty observed at clients' weaker dealer relationships is consistent with asymmetric information (and associated adverse selection risk) being higher for these trades.

Panel B of Table 7 presents the results for the case when we allow the size penalty to be heterogeneous between the dealers' strongest and weakest client relationships. We find that the size penalty becomes statistically insignificant for trades that the dealer conducts with its most important client. This is consistent with asymmetric information being negligible for these special relationships.[38]

3.5 The size penalty in government versus corporate bond markets

In this section, we assess potential variations in the size penalty between the corporate and government bond markets. Our ZEN data set, encompassing nearly all secondary market trades in U.K. corporate bonds, uniquely enables cross-market analysis. This feature facilitates the identification of clients engaged in active trading across both U.K.

[38] Note that the results are similar (though the heterogeneity is quantitatively weaker) when we consider the client's two most important dealers (Table G.14 in Internet Appendix) or when we consider the dealer's two most important clients (Table G.15 in Internet Appendix).

Table 7
Trading costs and trade size: The role of relationship strength

	(1)	(2)	(3)	(4)
A. Clients' weak versus strong (top-1) dealer relationship				
Clients' weaker dealer relationships				
Trade size	0.138***	0.153***	0.177***	0.202***
	(4.05)	(4.84)	(5.97)	(7.06)
Client's strongest (top-1) dealer relationships				
Trade size	0.050	0.088**	0.071*	0.073*
	(1.09)	(2.42)	(1.83)	(1.80)
p-values, eq. of coeff.	.036	.031	.003	.001
N	1,177,267	1,177,267	1,169,789	1,169,163
R^2	.107	.107	.198	.203
B. Dealers' weak versus strong (top-1) client relationship				
Dealers' weaker client relationships				
Trade size	0.116***	0.140***	0.150***	0.173***
	(3.28)	(4.55)	(4.76)	(5.71)
Dealers' strongest (top-1) client relationships				
Trade size	-0.026	0.009	0.054	0.054
	(-0.39)	(0.14)	(0.49)	(0.49)
p-values, eq. of coeff.	.048	.049	.415	.310
N	1,166,020	1,166,020	1,156,271	1,155,658
R^2	.084	.084	.165	.170
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	No	No
Dealer#ClientType FE	No	Yes	No	No
Day#Dealer#ClientType FE	No	No	Yes	Yes
Month#Client FE	No	No	Yes	Yes
Client#Dealer FE	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. The p-values correspond to the testing for the equality of coefficients. * $p < .1$; ** $p < .05$; *** $p < .01$.

bond markets.[39] Identifying a common set of clients is crucial for a cross-market comparison of the size penalty, because the client composition itself can be endogenously determined by the yields, riskiness, liquidity

[39] In a similar spirit, Appendix Section F compares the cost-size relation across the U.K. gilt market and the U.S. treasury market. In addition, we also revisit the cost-size relation during the COVID-19 period (Appendix Section E), which additionally provides an ideal setting for performing a cross-check using another, more recent data sample (2018–2020). We continue to find significant size penalty in this more recent data set, and additionally show evidence on an increased size penalty during the financial market turmoil in March 2020, which is particularly strong for the group of more sophisticated clients.

and opaqueness of the market in question (Dow 2004). We mitigate the selection issue by restricting the sample to clients who have a nontrivial presence in both types of bond markets.[40]

Possible differences in size penalty across the two markets could be driven by explanations related to information, liquidity or dealers' inventory costs (Hotchkiss and Jostova 2017; Friewald and Nagler 2019; Bessembinder, Spatt, and Venkataraman 2020). The identification assumption underlying our triple differences approach is that the liquidity and inventory mechanisms should generate a differential degree of size penalty in corporate bonds vis-à-vis in government bonds, *regardless of client type*. Therefore, if we find that the size penalty is larger for corporate bonds and, importantly, this increase is significantly larger for sophisticated clients than less sophisticated clients, then we can plausibly argue that information-based factors likely play a role in determining the size penalty over and above what is captured by liquidity and inventory channels.

To proceed, we sort all transactions into two groups $l = \{\text{GovBond, CorBond}\}$, based on whether the given trade occurred in the government bond or corporate bond market. We then estimate a modified version of our baseline regression 4, as follows:

$$Cost_v = \sum_{w=1}^{2}\sum_{z=1}^{2} \eta_{w,z} \times \mathbf{1}[j \in l_z, i \in g_w] \times Size_v + FE + \varepsilon_v, \qquad (12)$$

where $\mathbf{1}[j \in l_z, i \in g_w]$ is an indicator function equal to one when a bond j belongs to group z and client i belongs to group w, and is equal to zero otherwise; the term FE includes various combinations of fixed effects discussed above.

The results are shown in Table 8. We find that the size penalty is significantly larger in corporate bonds than in government bonds, but only for the sophisticated group of clients. While the most conservative specification (column 4) shows that the point estimate on the size penalty is about 0.225 bps larger (0.35 vs. 0.125) for less sophisticated clients, it is larger by about 0.67 bps (0.856 vs. 0.186) for the more sophisticated client groups. For the latter group, the statistical significance of these differences is particularly strong, with the corresponding p-values being less than .01. As shown by Tables G.8–G.9 in the Internet Appendix, we obtain similar results when using four alternative definitions of trading costs. Moreover, including all clients

[40] Specifically, we require that any client in this subsample must generate at least 15% of their volume in both markets. For example, this means that we omit most foreign central banks from this exercise, as they have little presence in the U.K. corporate bond markets and trade almost exclusively in government bonds. Some asset managers specialize in trading in either the government bond or the corporate bond market so they are also excluded from the sample.

Table 8
Trading costs and trade size: Government versus corporate bonds markets

	(1)	(2)	(3)	(4)
Less sophisticated clients				
Trade size#GovernmentBonds	0.078	0.093**	0.122***	0.125***
	(1.42)	(2.01)	(2.69)	(2.74)
Trade Size#CorporateBonds	0.310	0.316	0.332	0.350*
	(1.59)	(1.59)	(1.64)	(1.71)
p-values, eq. of coeff.	.149	.192	.232	.208
More sophisticated clients				
Trade Size#GovernmentBonds	0.152***	0.177***	0.186***	0.186***
	(3.92)	(4.70)	(5.16)	(5.11)
Trade Size#CorporateBonds	0.774***	0.749***	0.800***	0.856***
	(3.89)	(3.81)	(4.07)	(4.34)
p-values, eq. of coeff.	0.001	0.002	0.001	0.000
N	1,171,526	1,165,674	1,165,359	1,164,790
R^2	.349	.426	.430	.433
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes
Client#Market FE	Yes	Yes	Yes	No
Dealer#Market#ClientType FE	Yes	Yes	Yes	No
Day#Dealer#ClientType FE	No	Yes	Yes	Yes
Month#Client FE	No	Yes	Yes	Yes
Client#Dealer	No	No	Yes	No
Client#Dealer#Market FE	No	No	No	Yes

This table regresses trading costs (measured in bp) on trade size (measured as the logarithm of the nominal size of the trade in £s) interacted with an indicator variable taking value 2 (1) if the trade takes place in the corporate (government) bond market. The regression also includes various fixed effects. The upper (lower) panel shows the results for less (more) sophisticated clients. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. The p-values correspond to the testing for the equality of coefficients, within a given client type. * $p<.1$; ** $p<.05$; *** $p<.01$.

in the analysis (regardless of their relative trading volume in the two markets) lead to similar findings as well, as shown in Table G.10 in the Internet Appendix.

4. Conclusion

To conclude, our paper revisited the empirical relation between trade size and trading costs, one of the long-standing questions in the literature on financial markets. In our empirical design, we were able to observe client identity as well as their trading activities in U.K. government and corporate bond markets (as well as across U.K. and U.S. government bond markets). These unique features of our empirical design allowed us to reconcile some of the tension in the vast literature on the cost-size relation. Our results reveal that controlling for traders' identity is crucial for understanding the drivers of trading costs in nonanonymous over-the-counter markets. In addition, combining this client-level variation with variations in client-type, macroeconomic news and bond markets highlights the different forces that drive the cost-size relation.

There are a number of interesting avenues for future research. First, revisiting the cost-size relation around government bond auctions would be an interesting empirical extension of our framework. This could shed some interesting light on both supply effects (Lou, Yan, and Zhang 2013; Fleming and Liu 2017) and informational effects (Hortacsu and Kastl 2012; Boyarchenko, Lucca, and Veldkamp 2021) around primary issuances that can have important policy implications as well. Second, analyzing the size discount and size penalty in a structural framework could help provide a sharper characterization of the drivers of the cost-size relation and to quantify the relative importance of the different channels. Estimating the theoretical model presented in Internet Appendix A in the spirit of Odders-White and Ready (2008) could be a step into that direction. Third, one could consider the aggregate implications of our empirical analysis. For example, one could estimate how time-series variation in either the size discount or the size penalty is related to variation in aggregate bid-ask spreads and yields in government and corporate bond markets. This could tighten the link between our analysis and the literature on the term structure of interest rates.

Code Availability: The replication code is available in the Harvard Dataverse.

References

AFME. 2018. Government bond data report. Report, Association for Financial Markets in Europe.

Arnold, G. E., and M. E. Rhodes. 2021. Information sensitivity of corporate bonds: Evidence from the covid-19 crisis. *Finance Research Letters* 42:101911.

Ashtari, O. T., R. Guimaraes, G. Pinter, and J.-C. Wijnandts. 2023. The liquidity state-dependence of monetary policy transmission. Mimeo, Bank of England.

Benos, E., and F. Zikes. 2018. Funding constraints and liquidity in two-tiered OTC markets. *Journal of Financial Markets* 39:24–43.

Bernile, G., J. Hu, and Y. Tang. 2016. Can information be locked up? Informed trading ahead of macro-news announcements. *Journal of Financial Economics* 121:496–520.

Bessembinder, H., S. Jacobsen, W. Maxwell, and K. Venkataraman. 2018. Capital commitment and illiquidity in corporate bonds. *Journal of Finance* 73:1615–61.

Bessembinder, H., W. Maxwell, and K. Venkataraman. 2006. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics* 82:251–88.

Bessembinder, H., C. Spatt, and K. Venkataraman. 2020. A survey of the microstructure of fixed-income markets. *Journal of Financial and Quantitative Analysis* 55:1–45.

Biais, B. 1993. Price information and equilibrium liquidity in fragmented and centralized markets. *Journal of Finance* 48:157–85.

Biais, B., and R. Green. 2019. The microstructure of the bond market in the 20th century. *Review of Economic Dynamics* 33:250–71.

Boyarchenko, N., D. O. Lucca, and L. Veldkamp. 2021. Taking orders and taking notes: Dealer information sharing in treasury auctions. *Journal of Political Economy* 129:607–45.

Brancati, E., and M. Macchiavelli. 2019. The information sensitivity of debt in good and bad times. *Journal of Financial Economics* 133:99–112.

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* 59:2623–54.

Campbell, J., T. Ramadorai, and A. Schwartz. 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics* 92:66–91.

Chen, L., D. A. Lesmond, and J. Wei. 2007. Corporate yield spreads and bond liquidity. *Journal of Finance* 62:119–49.

Czech, R., S. Huang, D. Lou, and T. Wang. 2021. Informed trading in government bond markets. *Journal of Financial Economics* 142:1253–74.

Czech, R., and G. Pinter. 2020. Informed trading and the dynamics of client-dealer connections in corporate bond markets. Working Paper, Bank of England.

Di Maggio, M., F. Franzoni, A. Kermani, and C. Sommavilla. 2019. The relevance of broker networks for information diffusion in the stock market. *Journal of Financial Economics* 134:419–46.

Di Maggio, M., A. Kermani, and Z. Song. 2017. The value of trading relations in turbulent times. *Journal of Financial Economics* 124:266–84.

DMO. 2011. A guide to the roles of the DMO and Primary Dealers in the UK government bond market. Working Paper, Debt Management Office.

———. 2021. A guide to the roles of the DMO and Primary Dealers in the UK government bond market. Working Paper, Debt Management Office.

Dow, J. 2004. Is liquidity self-fulfilling? *Journal of Business* 77:895–908.

Du, B., S. Fung, and R. Loveland. 2018. The informational role of options markets: Evidence from FOMC announcements. *Journal of Banking & Finance* 92:237–56.

Du, S., and H. Zhu. 2017. Bilateral trading in divisible double auctions. *Journal of Economic Theory* 167:285–311.

Duffie, D. 2020. Still the world's safe haven? Redesigning the U.S. Treasury market after the COVID-19 crisis. Working Paper, Brookings.

Duffie, D., N. Gârleanu, and L. H. Pedersen. 2005. Over-the-counter markets. *Econometrica* 73:1815–47.

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

Edwards, A., L. E. Harris, and M. S. Piwowar. 2007. Corporate bond market transaction costs and transparency. *Journal of Finance* 62:1421–51.

Eguren-Martin, F., and N. McLaren. 2015. How much do UK market interest rates respond to macroeconomic data news? *Bank of England Quarterly Bulletin* 55:259–72.

Falato, A., I. Goldstein, and A. Hortacsu. 2021. Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets. *Journal of Monetary Economics* 123:35–52.

Farboodi, M., and L. Veldkamp. 2020. Long-run growth of financial data technology. *American Economic Review* 110:2485–523.

Feldhutter, P. 2012. The same bond at different prices: Identifying search frictions and selling pressures. *Review of Financial Studies* 25:1155–206.

Fleming, M. J. 1997. The round-the-clock market for U.S. Treasury securities. *Economic Policy Review* 3:9–32.

Fleming, M. J., and W. Liu. 2017. Intraday pricing and liquidity effects of U.S. Treasury auctions. Working Paper, Federal Reserve Bank of New York.

Fleming, M. J., B. Mizrach, and G. Nguyen. 2018. The microstructure of a U.S. Treasury ECN: The BrokerTec platform. *Journal of Financial Markets* 40:2–22.

Friewald, N., and F. Nagler. 2019. Over-the-counter market frictions and yield spread changes. *Journal of Finance* 74:3217–57.

Gallup, J. L. 2019. Added-variable plots with confidence intervals. *Stata Journal* 19:598–614.

Glosten, L. R. 1989. Insider trading, liquidity, and the role of the monopolist specialist. *Journal of Business* 62:211–35.

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* 14:71–100.

Goldstein, M. A., E. S. Hotchkiss, and E. R. Sirri. 2007. Transparency and liquidity: A controlled experiment on corporate bonds. *Review of Financial Studies* 20:235–73.

Golosov, M., G. Lorenzoni, and A. Tsyvinski. 2014. Decentralized trading with private information. *Econometrica* 82:1055–91.

Gould, J. P., and R. E. Verrecchia. 1985. The information content of specialist pricing. *Journal of Political Economy* 93:66–83.

Green, R., B. Hollifield, and N. Schurhoff. 2007a. Dealer intermediation and price behavior in the aftermarket for new bond issues. *Journal of Financial Economics* 86:643–82.

———. 2007b. Financial intermediation and the costs of trading in an opaque market. *Review of Financial Studies* 20:275–314.

Green, T. C. 2004. Economic news and the impact of trading on bond prices. *Journal of Finance* 59:1201–34.

Haddad, V., A. Moreira, and T. Muir. 2021. When selling becomes viral: Disruptions in debt markets in the COVID-19 crisis and the Fed's response. *Review of Financial Studies* 34:5309–51.

Harris, L. E., and M. S. Piwowar. 2006. Secondary trading costs in the municipal bond market. *Journal of Finance* 61:1361–97.

Hauser, A. 2020. Seven Moments in Spring: Covid-19, financial markets and the Bank of England's balance sheet operations. Speech, Bank of England.

He, Z., S. Nagel, and Z. Song. 2022. Treasury inconvenience yields during the COVID-19 crisis. *Journal of Financial Economics* 143:57–79.

Hendershott, T., R. Kozhan, and V. Raman. 2020. Short selling and price discovery in corporate bonds. *Journal of Financial and Quantitative Analysis* 55:77–115.

Hendershott, T., D. Li, D. Livdan, and N. Schürhoff. 2020. Relationship trading in over-the-counter markets. *Journal of Finance* 75:683–734.

Hendershott, T., and A. Madhavan. 2015. Click or call? auction versus search in the over-the-counter market. *Journal of Finance* 70:419–47.

Ho, T., and H. R. Stoll. 1981. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9:47–73.

Hoel, M. 1987. Bargaining games with a random sequence of who makes the offers. *Economics Letters* 24:5–9.

Hollifield, B., A. Neklyudov, and C. Spatt. 2017. Bid-ask spreads, trading networks, and the pricing of securitizations. *Review of Financial Studies* 30:3048–85.

Hortacsu, A., and J. Kastl. 2012. Valuing dealers' informational advantage: A study of Canadian Treasury auctions. *Econometrica* 80:2511–42.

Hotchkiss, E., and G. Jostova. 2017. Determinants of corporate bond trading: A comprehensive analysis. *Quarterly Journal of Finance* 7:1–30.

Ivanov, P., A. Orlov, and M. Schihl. 2021. Bond liquidity and dealer inventories: Insights from US and European regulatory data. Occasional Paper 52, Financial Conduct Authority.

Kargar, M., B. Lester, D. Lindsay, S. Liu, P.-O. Weill, and D. Zuniga. 2021. Corporate bond liquidity during the COVID-19 crisis. *Review of Financial Studies* 34:5352–401.

Klingler, S., and S. Sundaresan. 2019. An explanation of negative swap spreads: Demand for duration from underfunded pension plans. *Journal of Finance* 74:675–710.

Kondor, P., and G. Pinter. 2022. Clients' connections: Measuring the role of private information in decentralized markets. *Journal of Finance* 77:505–44.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–35.

Kyle, A. S., and A. A. Obizhaeva. 2016. Market microstructure invariance: Empirical hypotheses. *Econometrica* 84:1345–404.

Lagos, R., and G. Rocheteau. 2009. Liquidity in asset markets with search frictions. *Econometrica* 77:403–26.

Leach, J. C., and A. Madhavan. 1993. Price experimentation and security market structure. *Review of Financial Studies* 6:375–404.

Lester, B., A. Shourideh, V. Venkateswaran, and A. Zetlin-Jones. 2023. Market-making with search and information frictions. *Journal of Economic Theory* 212:105714.

Li, D., and N. Schürhoff. 2019. Dealer networks. *Journal of Finance* 74:91–144.

Lou, D., H. Yan, and J. Zhang. 2013. Anticipated and repeated shocks in liquid markets. *Review of Financial Studies* 26:1891–912.

Ma, Y., K. Xiao, and Y. Zeng. 2022. Mutual fund liquidity transformation and reverse flight to liquidity. *Review of Financial Studies* 35:4674–711.

Odders-White, E. R., and M. J. Ready. 2008. The probability and magnitude of information events. *Journal of Financial Economics* 87:227–48.

O'Hara, M., Y. Wang, and X. Zhou. 2018. The execution quality of corporate bonds. *Journal of Financial Economics* 130:308–26.

O'Hara, M., and X. A. Zhou. 2021. Anatomy of a liquidity crisis: Corporate bonds in the COVID-19 crisis. *Journal of Financial Economics* 142:46–68.

Osler, C. L., A. Mende, and L. Menkhoff. 2011. Price discovery in currency markets. *Journal of International Money and Finance* 30:1696–718.

Pasquariello, P., and C. Vega. 2007. Informed and strategic order flow in the bond markets. *Review of Financial Studies* 20:1975–2019.

Pinter, G. 2023. An anatomy of the 2022 gilt market crisis. Working Paper, Bank of England.

Pinter, G., and S. Uslu. 2021. Comparing search and intermediation frictions across markets. Mimeo.

Pinter, G., C. Wang, and J. Zou. 2020. Information chasing and adverse selection in over-the-counter markets. Working Paper, Bank of England.

Ramadorai, T. 2008. What determines transaction costs in foreign exchange markets? *International Journal of Finance & Economics* 13:14–25.

Rubinstein, A. 1982. Perfect equilibrium in a bargaining model. *Econometrica* 50:97–109.

Schultz, P. 2001. Corporate bond trading costs: A peek behind the curtain. *Journal of Finance* 56:677–98.

Smith, B. F., D. A. S. Turnbull, and R. W. White. 2001. Upstairs market for principal and agency trades: Analysis of adverse information and price effects. *Journal of Finance* 56:1723–46.

Subrahmanyam, A. 1991. Risk aversion, market liquidity, and price efficiency. *Review of Financial Studies* 4:416–41.

Swanson, E. T., and J. C. Williams. 2014. Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review* 104:3154–85.

US-Treasury. 2021. Recent disruptions and potential reforms in the U.S. Treasury market: A staff progress report. Progress Report, U.S. Department of the Treasury.

Vayanos, D., and J.-L. Vila. 2021. A preferred-habitat model of the term structure of interest rates. *Econometrica* 89:77–112.

Viswanathan, S., and J. J. D. Wang. 2004. Inter-dealer trading in financial markets. *Journal of Business* 77:987–1040.

Vives, X. 2011. Strategic supply function competition with private information. *Econometrica* 79:1919–66.

Internet Appendix

A. A Theoretical Model of Bilateral Trading

A.1 Overview

In this section, we present a bilateral trading model with bargaining and information asymmetry to rationalise the co-existence of the size discount and penalty. The model builds on the standard CARA-normal setting with strategic traders as in Kyle (1985) and risk aversion Subrahmanyam (1991) with the addition of a liquidity component. Importantly, we show how one can introduce into this framework of asymmetric information a simple bargaining game (building on Rubinstein (1982) and Hoel (1987)) which can generate both the size discount and the size penalty. This model could be a building block for analyzing strategic bilateral trading under asymmetric information in decentralised markets.

Combining information asymmetry and bargaining power is, in general, a hard theoretical problem. In most papers, prices are either monopolistically (Gould and Verrecchia 1985; Glosten 1989) or competitively (Glosten and Milgrom 1985; Kyle 1985) offered by an uninformed party to an informed party. This simple price-setting mechanism avoids signalling and screening, at the expense of not being able to achieve surplus splitting between the two parties.[41] We circumvent the technical difficulty of bargaining under asymmetric information, by showing that a trader with both liquidity and informational trading motives would perfectly reveal her joint trading motive through her order size in a linear-pricing equilibrium without bargaining delay. Our model jointly predicts within-client size penalty and across-client size discount. Each of the two channels can be conveniently shut down, in which case the model reduces to a standard model with remaining feature explaining either size penalty or size discount.

A.2 The Model

A client seeks to trade a risky asset with a dealer. The value of the asset, v, has an unconditional distribution $N\left(0, \frac{1}{\tau_v}\right)$. The client observes a noisy signal, $s = v + \varepsilon$, before the trade. The noise ε is normally distributed, $N\left(0, \frac{1}{\tau_\varepsilon}\right)$. The client has an initial asset position of x, following a normal distribution $N\left(0, \frac{1}{\tau_x}\right)$. The dealer does not observe the client's signal, s, or her initial position, x. The random variables v, ε and x are jointly independent. The client has a CARA utility

[41] In Lester et al. (2023), the price can be either competitive or monopolistic with an exogenous probability. While not focusing on surplus splitting, Du and Zhu (2017) studies a bilateral double-auction model with information asymmetry in which the price is set by a double auction between two traders.

function with risk aversion γ. The initial risky asset position x gives the client a liquidity motive to trade, while the signal s gives her an informational motive. The client's preference and information follow the standard CARA-normal setting with strategic traders as in Kyle (1985) and risk aversion as in Subrahmanyam (1991), with the addition of a liquidity component, x, added to her trading motive. The dealer is risk neutral.

A trade is conducted as follows. First, the client requests to buy q unit of the asset from the dealer ($q<0$ means that the client requests to sell). Then, to negotiate a price, the client and the dealer engage in an infinite-horizon bargaining game with discount rate δ and a random sequence of who makes the offers. In the bargaining game, the dealer moves first to offer a price in response to the client's request. The bargaining game concludes if the client accepts the price offer, and otherwise continues to the next stage. In each subsequent stage, the client is selected to offer a price with probability η to the dealer, and the other way round with probability $1-\eta$ as in Hoel (1987), which is adapted from the alternating offer game of Rubinstein (1982).

We consider a Perfect Bayesian Equilibrium (PBE) in which the agreed-upon price p is linear in size, $p(q)=a+\lambda q$, and the first price offer is immediately accepted. Such an equilibrium is said to be a *linear-pricing PBE without bargaining delay.* The following theorem summarizes the equilibrium of the model.

Theorem A.1. When $\left(2\frac{\eta}{1+\delta}-1\right)\gamma^2\tau_v^2-\tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon)>0$, there exists a unique linear-pricing PBE without bargaining delay. On the equilibrium path, the client submits an order $q_\eta^*(s,x)$, the dealer offers $p_\eta^*(q)=\lambda_\eta^* q$, and the client immediately accepts the price offer, where

$$\lambda_\eta^*=\frac{\gamma}{2}\frac{\left(1+\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon)}\right)+\frac{\eta}{1+\delta}\left(1-\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon)}\right)}{\left(2\frac{\eta}{1+\delta}-1\right)\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x}-(\tau_v+\tau_\varepsilon)}$$

$$q_\eta^*(s,x)=\frac{\gamma}{2\lambda^*(\tau_v+\tau_\varepsilon)+\gamma}\left(\frac{\tau_\varepsilon}{\gamma}s-x\right).$$

We establish Theorem A.1 by solving for a linear-pricing PBE without bargaining delay, with the key steps detailed in the following proof.

Proof. Given each signal s and initial position x, the client's expected gain from trading q units of the asset is:

$$\frac{\tau_\varepsilon}{\tau_\varepsilon+\tau_v}s(q+x)-pq-\frac{\gamma}{2}\frac{1}{\tau_\varepsilon+\tau_v}(q+x)^2-\left[\frac{\tau_\varepsilon}{\tau_\varepsilon+\tau_v}sx-\frac{\gamma}{2}\frac{1}{\tau_\varepsilon+\tau_v}x^2\right]\equiv(p_0(q,s,x)-p)q,$$

where $p_0(q,s,x)$ is the client's reservation price:

$$p_0(q,s,x) = \gamma \frac{1}{\tau_\varepsilon + \tau_v}\left(\frac{\tau_\varepsilon}{\gamma}s - x\right) - \frac{\gamma}{2}\frac{1}{\tau_\varepsilon+\tau_v}q. \tag{A.1}$$

Anticipating an equilibrium price $p(q)$ without bargaining delay, the client chooses size q to maximize her expected trading gain. Her first order condition is $p_0(q,s,x) - p(q) + \left(\frac{\partial p_0(q,s,x)}{\partial q} - \frac{\partial p(q)}{\partial q}\right)q = 0$ which, together with a linear equilibrium price function $p(q) = a + \lambda q$, can be used to obtained the client's optimal demand q:

$$q = \frac{\gamma}{2\lambda(\tau_\varepsilon+\tau_v)+\gamma}\left(\frac{\tau_\varepsilon}{\gamma}s - x\right) - \frac{\tau_\varepsilon+\tau_v}{2\lambda(\tau_\varepsilon+\tau_v)+\gamma}a. \tag{A.2}$$

With this optimal choice q, the client's expected gain $(p_0(q,s,x)-p)q$ from trading can be written as:

$$(p_0(q,s,x)-p)q = \left(\frac{2\lambda(\tau_\varepsilon+\tau_v)+\gamma}{\tau_\varepsilon+\tau_v} - \frac{\gamma}{2}\frac{1}{\tau_\varepsilon+\tau_v} + \lambda\right)q^2, \tag{A.3}$$

which shows that the client's expected gain from trading depends on her signal s and initial endowment x only through her requested size q. This equilibrium property renders the ensuing bargaining game one with complete information. Therefore, the solution of Hoel (1987) applies, whereby the dealer offers price $p(q)$ which is immediately accepted by the client:

$$p(q) = \frac{\eta}{1+\delta}p_1(q) + \left(1 - \frac{\eta}{1+\delta}\right)p_0(q), \tag{A.4}$$

where $p_1(q)$ is given by:

$$p_1(q) = \mathbb{E}[v|q] = \frac{2\lambda(\tau_v+\tau_\varepsilon)+\gamma}{\tau_v+\tau_\varepsilon+\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x}}q + \frac{\tau_v+\tau_\varepsilon}{\tau_v+\tau_\varepsilon+\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x}}a. \tag{A.5}$$

In A.5, $p_1(q)$ can be viewed as the competitive price, which gives the dealer zero expected profit.[42] In equilibrium, the dealer's price offer is anticipated by the client. Therefore, substituting A.1 and A.5 into A.4, and matching coefficients yields $a = 0$, while rearranging gives the solution for λ:

$$\lambda = \frac{\gamma}{2}\frac{\left(1+\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon)}\right) + \frac{\eta}{1+\delta}\left(1 - \frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon)}\right)}{\left(2\frac{\eta}{1+\delta}-1\right)\frac{\gamma^2\tau_v^2}{\tau_\varepsilon\tau_x} - (\tau_v+\tau_\varepsilon)}. \tag{A.6}$$

To verify that the size choice q in A.2 given by the client's FOC indeed maximizes the client's expected gain, we check the client's second order condition $2\lambda(\tau_v+\tau_\varepsilon)+\gamma > 0$, which is equivalent to $[2\eta/(1+\delta)-1]\gamma^2\tau_v^2 - \tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon) > 0$. ■

[42] To derive the conditional expectation $\mathbb{E}[v|q]$, we use the projection theorem. For that, note that the covariance between the asset value and trade size is $Cov[v,q] = (\tau_\varepsilon/\tau_v)/[2\lambda(\tau_\varepsilon+\tau_v)+\gamma]$, and the variance of the trade size is $Var[q] = [\gamma/(2\lambda(\tau_\varepsilon+\tau_v)+\gamma)]^2\left[(\tau_\varepsilon/\gamma)^2/\tau_\varepsilon + (\tau_\varepsilon/\gamma)^2/\tau_v + 1/\tau_x\right]$.

The key property allowing for a tractable bargaining solution is that the client's expected gain from trading A.3 (and thereby the client's reservation price p_0) does not directly depend on s and x but only through the requested size q. Even though the client's liquidity motive x and informational motive s are not uniquely determined by her requested size q, the aggregate motive is. This nice equilibrium property renders the ensuing bargaining game one with complete information, so that the solution of Hoel (1987) applies.

The model simultaneously explains within-client size penalty and across-client size discount:

Proposition A.1. (Within-client size penalty) In the linear-pricing PBE without delay, the trading cost $|p(q)| = \lambda|q|$ of a given client with bargaining power η is linearly increasing with her requested size $|q|$.

Next, we compare two clients with heterogeneous bargaining powers $\eta < \eta'$ to generate size discount.

Proposition A.2. (Across-client size discount) In the linear-pricing PBE without delay, a client with higher bargaining power η has a larger average trade size $\mathbb{E}|q|$, while facing a lower average trading cost $\mathbb{E}|p|$.

The above result follows immediately from the fact the average trade size $\mathbb{E}|q|$ is decreasing in $\lambda(\eta)$, while the average trading cost $\mathbb{E}|p|$ is increasing in $\lambda(\eta)$. On the other hand, $\lambda(\eta)$ is decreasing in bargaining power η.

A.3 Discussion

To highlight the novel features of the model, we perform simple comparative statics to explore the relationship between bargaining power and the size penalty as shown in Figure A.1.

The Figure plots the values of λ (vertical axis), as computed by A.6, for different values of τ_ε (horizontal axis). We trace out the relationship for three different values $\eta = \{0.7, 0.75, 0.8\}$, as shown by the dashed, dotted and solid lines, respectively. All three lines are monotonically increasing: the larger the precision of the signal (τ_ε) the larger the size penalty becomes. The intuition is similar to Kyle (1985): the more precise the informed client's signal, the more informative her trade becomes, which makes the dealer revise the price more aggressively when considering trading a given quantity during the bargaining process. Moreover, one possible way to connect our client-type variation in the data with the model is to regard more sophisticated clients as having higher values of τ_ε, so that this simple theoretical framework could

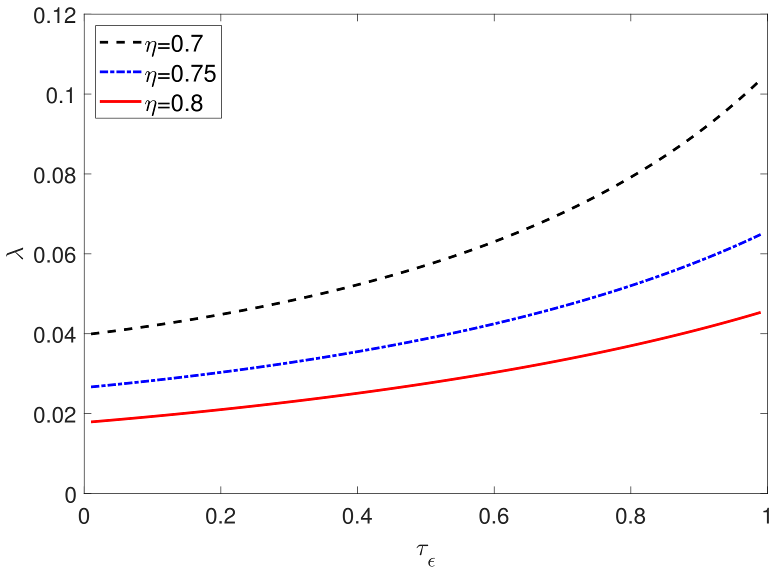


Figure A.1
Bargaining Power, Signal Precision and Size Penalty
This figure plots the value of λ (computed by A.6) for different values of $\tau_\epsilon = [0.01 : 0.01 : 0.99]$ and of $\eta = \{0.7, 0.75, 0.8\}$. The rest of the parameters are $\gamma = 0.2$, $\tau_v = 2$, $\tau_x = 0.1$, $\delta = 0.01$ so that the second order condition, $[2\eta/(1+\delta)-1]\gamma^2\tau_v^2 - \tau_\varepsilon\tau_x(\tau_v+\tau_\varepsilon) > 0$, is always satisfied.

rationalise why more sophisticated clients face a higher size penalty in the data than less sophisticated clients.

Furthermore, the τ_ε-λ relation largely depends on the client's bargaining power, η, as shown by the different lines in Figure A.1: lowering η is associated with an upward shift in the curve, which means that the size penalty becomes stronger as the client's bargaining power becomes weaker. In addition, we also find evidence for increased convexity in the τ_ε-λ relation as the client's bargaining power decreases, which means that the size penalty becomes increasingly sensitive to the precision of private information. This highlights some of the rich interactions between informational and trading frictions that this simple model can generate, and which goes beyond the regression design of our empirical model. Future work could undertake a structural estimation of a version of this model (in the spirit of Odders-White and Ready (2008)) in order to explore the model's predictions further.

B. Data Construction

The main data source for our regression results is the ZEN data set which was the United Kingdom’s transaction reporting system administered by the Financial Conduct Authority. A detailed description of transaction reporting in ZEN can be found here:
https://www.fca.org.uk/publication/finalised-guidance/fg15-03.pdf.
Our sample period covers 2011m8-2017m12. There is no public version of ZEN, which is why this data set has been used relatively rarely in the Academic literature (recent exceptions include Benos and Zikes (2018), Czech et al. (2021), Pinter and Uslu (2021) and Kondor and Pinter (2022)). The structure of the ZEN data set is similar to the TRACE data set that is often used to study the U.S. corporate bond market. A notable difference compared to TRACE is that nearly all trade reports in ZEN include counterparty identities. (Ivanov, Orlov, and Schihl (2021) provides a recent comparison between the ZEN data set and the TRACE data set, using a common set of corporate bonds traded in both the United Kingdom and United States).

The ZEN data set includes all secondary market trades, where at least one of the counterparties is an FCA-regulated entity. We drop duplicate trade reports as well as trade reports with missing client identifiers. We also exclude trades with interdealer brokers (IDBs), dealer-to-dealer trades, and client-to-client trades as well as trades of less than £1,000 in par value, and remove trades with erroneous price entries. We include both principal and agency trades in our sample. In our baseline sample, we end up with around 1.27 million observations for government bond trades, covering 57 government bonds and around 600 clients. About half of these clients can be classified as hedge funds and asset managers, and the remaining half comprises of pension funds, insurance companies, commercial banks, international organisations and others. Our sample of corporate bond trades (covering the sample period) includes around 1.23 million observations and around 2,800 corporate bonds.

Our sample of U.S. Treasury securities also come from the ZEN data set. Similar to our treatment of the U.K. data, we exclude trades with interdealer brokers (IDBs), dealer-to-dealer trades, and client-to-client trades as well as trades of less than $1,000 in par value, and remove trades with erroneous price entries. We search for the same set of dealers and clients (as in our sample of U.K. gilts) and end up with around 1.3 million transactions for around 950 U.S. Treasuries. We find that about 75% of our sample of U.K. gilt market clients trade Treasuries in our sample of U.S. trades.

C. Controlling for Observable Characteristics

Our baseline regression Table 2 includes various combination of fixed effects as control variables. While the interpretation of the multiple fixed effects is motivated in detail in Section 2.1, we include here an alternative version of Table 2 of the paper in order to increase the transparency of our empirical model. Specifically, we estimate a version of regression 2 whereby we explicitly control for observable characteristics instead of using fixed effects. Column 1 of Table C.1 includes trade volume, number of transactions and price dispersion (Pinter and Uslu 2021) in place of bond-day fixed effects. These variables aim to control for liquidity conditions that may vary from one day to the next, possibly heterogeneously across bonds.

Table C.1
Trading Costs and Trade Size in Government Bond Markets: Including Observable Control Variables

	(1)	(2)	(3)	(4)	(5)
Trade Size	-0.240***	0.073**	0.097***	0.133***	0.167***
	(-4.00)	(2.04)	(2.78)	(4.19)	(5.61)
Trade Volume (Day-Bond)	0.072	0.060	0.052	0.070*	0.076*
	(1.57)	(1.38)	(1.21)	(1.73)	(1.93)
Transaction Nr. (Day-Bond)	-0.062	-0.059	-0.108	-0.071	-0.099
	(-0.59)	(-0.57)	(-1.06)	(-0.70)	(-0.97)
Price Dispersion (Day-Bond)	0.086	0.180*	0.221**	0.261**	0.267**
	(0.84)	(1.68)	(2.10)	(2.46)	(2.52)
Average Monthly Client Volume		-0.529***	-0.407***	-0.316***	-0.121*
		(-8.12)	(-6.37)	(-6.04)	(-1.95)
Monthly Client Volume			-0.176***	-0.109***	-0.200***
			(-5.54)	(-3.87)	(-4.47)
Daily Dealer Volume				-0.533***	-0.493***
				(-7.83)	(-7.38)
Total Client-Dealer Volume					-0.194**
					(-2.23)
Total Client-Dealer Tran. Nr.					0.416***
					(4.14)
N	1274238	1274238	1274238	1274238	1274238
R^2	0.001	0.002	0.003	0.004	0.004

This table regresses trading costs on trade size and various control variables (2). The cost measure is in basis points. The control variables are as follows: trade volume, number of transactions and price dispersion at the day-bond level (column 1); average monthly client volume (column 2); total monthly client volume (column 3); total daily dealer volume (column 4); total client-dealer trade volume as well as total number of transactions (column 5). All control variables are in logs. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

Column 2 of Table C.1 includes average monthly trade volume at the client level, which we use as a proxy for trader size. Analogous to the role played by client fixed effects in our baseline results, the inclusion of this trader size proxy makes the largest marginal change to the estimated coefficient on trade size by flipping the size and changing the estimate from -0.240 (column 1) to 0.073 (column 2). Column 3, 4, and 5 include additionally monthly client volume (analogous to client-month fixed effects in Table 2), daily dealer volume (analogous to dealer-day fixed effects) and total client-dealer volume as well as total number of transactions for each client-dealer pair (analogous to client-dealer fixed effects), respectively. Similar to our baseline results, the inclusion of these variable increases the estimated size penalty from 0.073 to 0.167.

Note however that the fixed effect specifications result in a substantially higher explanatory power, with an R^2-statistics of around 14% in the most conservative fixed effect model in comparison with an R^2 of less than 1% in the corresponding regression specification with observed control variables.

D. Dynamic Effects

An interesting question relates to how the size penalty behaves for the more and less sophisticated clients in the days leading up to the event and afterwards. To address this point, we estimate a dynamic version of Equation 4 i, whereby we take leads and lags of the dummy variable indicating whether the trading day coincides with large or small macroeconomic announcements. When estimating these models, we use the most conservative fixed effect specification corresponding to column 5 of Table 2.

Figure D.1 shows the coefficients on the size penalty around large macroeconomic news for the group of less sophisticated (left panel) and more sophisticated clients (right panel). There are no visible trend in size penalty for the group of less sophisticated clients. In contrast, as shown by the right panel of Figure D.1, the size penalty gradually increases as we get closer to the day of a big macroeconomic announcement, with the corresponding point estimate rising from around 0.16 and peaking at around 0.22 on the day of the announcement. These results suggest that the increased size penalty faced by more sophisticated traders reflect that dealers expect these clients' trades to become more informative as the macroeconomic news approaches.

E. COVID-19

We also study the cost-size relation during the COVID-19 episode in the United Kingdom. The spread of the COVID-19 pandemic in early

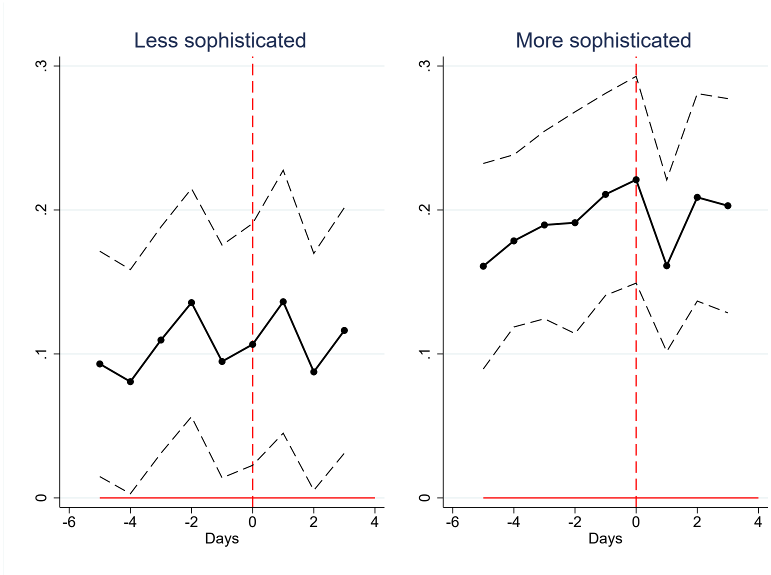


Figure D.1
The Dynamics of Size Penalty around Big Macroeconomic News
This figure plots the estimated coefficients from a variant of regression xx where we use leads and lags of the dummy variable indicating whether the trading day coincides with large or small macroeconomic announcements. The macroeconomic surprises are constructed following the high-frequency methodology of Eguren-Martin and McLaren (2015); Swanson and Williams (2014). The regression also includes various fixed effects described in column 5 of Table 2. The 90% confidence interval based on robust standard errors, using two-way clustering at the day and client level.

2020 presented a major shock to the global financial system, including bond markets. Investigating the behavior of the cost-size relation during this informationally intensive period provides an ideal opportunity to perform an out-of-sample robustness test of our baseline results. This is because more recent sample period (2018-2020) requires the use of a different data set compared to our baseline sample (2011-2017).

In addition, a better understanding of the functioning of government bond markets after COVID-19 is interesting on its own right and it is becoming ever more important for policy design (Duffie 2020; Hauser 2020). While a growing literature has analyzed the unfolding of the crisis in bond markets and the effect of subsequent central bank interventions, the majority of this literature focused on corporate bond markets in the United States (see, e.g., Falato, Goldstein, and Hortacsu 2021; Haddad, Moreira, and Muir 2021; Kargar et al. 2021; O'Hara and Zhou 2021; Ma, Xiao, and Zeng 2022) and the United Kingdom (Czech and Pinter

2020), and there has been little transaction-level evidence on the effect of COVID-19 in government bond markets.[43]

Table E.1
Trading Costs and Trade Size: During and Outside COVID-19

	(1)	(2)	(3)	(4)
Less Sophisticated Clients				
Trade Size#OutsideCOVID-19	0.261***	0.257***	0.251***	0.214***
	(4.20)	(4.43)	(4.26)	(3.56)
Trade Size#DuringCOVID-19	0.404*	0.301	0.274	0.273
	(1.81)	(1.62)	(1.43)	(1.57)
p-values, eq. of coeff.	0.525	0.808	0.903	0.737
More Sophisticated Clients				
Trade Size#OutsideCOVID-19	0.318***	0.294***	0.311***	0.329***
	(6.92)	(6.81)	(7.80)	(8.45)
Trade Size#DuringCOVID-19	0.581**	0.611***	0.635***	0.652***
	(2.22)	(2.86)	(3.12)	(3.40)
p-values, eq. of coeff.	0.271	0.121	0.096	0.080
N	1143362	1142116	1141464	1114966
R^2	0.146	0.194	0.203	0.262
Day#Bond FE	Yes	Yes	Yes	Yes
Dealer FE	Yes	No	No	No
Month#Client FE	Yes	Yes	Yes	No
Day#Dealer FE	No	Yes	Yes	Yes
Client#Dealer FE	No	No	Yes	No
Client#Dealer#Month FE	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients, within a given client type.

To carry out the analysis, we employ the MiFID II bond transaction data, which covers the period from January 2018 to July 2020.[44] Similar to the ZEN data, the MiFID II data provide detailed information (including counterparty identifiers) on transactions in the

[43] For a theoretical analysis of COVID-19 crisis in U.S. government bond markets, see He, Nagel, and Song (2022).

[44] The MiFID II reporting requirements became applicable on 3 January 2018. While ZEN is generally regarded as the predecessor of the MiFID II database, there are significant differences in the reporting requirements that prohibit a consistent merge of both data sets.

U.K. corporate bond market and give us almost full coverage of the client trade universe.

The following analysis serves two purposes. First, we check whether size penalty continues to hold in a different and more recent sample, and whether we continue to find a more pronounced effect for more sophisticated clients, thereby providing additional tests for Hypotheses 2-4 of our paper. Second, we explore how the size penalty might have changed during the unprecedented COVID-19 crisis period. To this end, we sort all the trades into two groups $c=\{c_1,c_2\}$ based on whether the trade occurred during March 2020 or outside this month. We then estimate the following regression:

$$Cost_v=\sum_{w=1}^{2}\sum_{z=1}^{2}\eta_{w,z}\times \mathbf{1}_t[t\in c_z,i\in g_w]\times Size_v+FE+\varepsilon_v, \qquad \text{(E.1)}$$

where $\mathbf{1}_t[t\in c_z,i\in g_w]$ is an indicator function equal to one when a trading day t belongs to group z and client i belongs to group w, and is equal to zero otherwise; the term FE includes various combinations of fixed effects discussed in the main text.

Table E.1 shows the results from estimating regression E.1, first for the group of less sophisticated (upper panel), and then for more sophisticated clients (lower panel). We continue to find that more sophisticated clients face a larger size penalty than less sophisticated clients during normal times, with the difference being about 0.1 bps (0.214 vs. 0.329) according to the most conservative specification in column (4). This difference is similar to our baseline based on the ZEN data for 2011-2017. Importantly, we find that the size penalty increases considerably during the COVID crisis and this increase is more pronounced for more sophisticated clients (0.652) compared to the other client type (0.273) – consistent with Hypothesis 4 and Table 4 in the main text.

We acknowledge however that these results could be driven by the fact that our group of more sophisticated clients mainly consists of asset managers who could have been under severe selling pressure as documented by the recent literature.[45] Therefore, we try to control for the possible contribution of selling pressure to the increase in the size penalty of more sophisticated clients during the COVID-19 crisis by adding an additional layer of differences to our research design. Specifically, we sort our group of more sophisticated clients into two subgroups based on the cumulative signed order flow in March 2020,

[45] See Falato, Goldstein, and Hortacsu (2021); Haddad, Moreira, and Muir (2021); Kargar et al. (2021); Ma, Xiao, and Zeng (2022); O'Hara and Zhou (2021) amongst others.

Table E.2
Trading Costs and Trade Size: During and Outside COVID-19

	(1)	(2)	(3)	(4)
More Sophisticated Clients Under More Selling Pressure				
Trade Size#OutsideCOVID-19	0.320***	0.296***	0.321***	0.339***
	(5.59)	(5.84)	(7.03)	(7.47)
Trade Size#DuringCOVID-19	0.605**	0.637***	0.647***	0.644***
	(2.23)	(2.99)	(3.15)	(3.05)
p-values, eq. of coeff.	0.226	0.098	0.098	0.138
N	483475	482910	482819	476606
R^2	0.153	0.212	0.217	0.261
More Sophisticated Clients Under Less Selling Pressure				
Trade Size#OutsideCOVID-19	0.348***	0.314***	0.313***	0.311***
	(6.66)	(5.73)	(5.89)	(5.93)
Trade Size#DuringCOVID-19	0.576	0.475	0.511*	0.664***
	(1.53)	(1.58)	(1.80)	(3.11)
p-values, eq. of coeff.	0.535	0.602	0.494	0.114
N	206443	205745	205629	200467
R^2	0.241	0.325	0.333	0.393
Day#Bond FE	Yes	Yes	Yes	Yes
Dealer FE	Yes	No	No	No
Month#Client FE	Yes	Yes	Yes	No
Day#Dealer FE	No	Yes	Yes	Yes
Client#Dealer FE	No	No	Yes	No
Client#Dealer#Month FE	No	No	No	Yes

This table regresses trading cost on trade size and various fixed effects. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p < .1$, ** $p < .05$, *** $p < .01$). The p-values correspond to the testing for the equality of coefficients, within a given client type.

and we then estimate our cost-size regression separately for these two subgroups.

Table E.2 shows the results. Based on the point estimates, we find that while the increase in size-penalty was large for clients under selling pressure (0.644 vs. 0.339), there is still a similarly larger increase in the size penalty for clients who were under less selling pressure (0.664 vs. 0.311).

Note also that the regressions include client-dealer-month fixed effects that aim to control for relationship effects during turbulent times (Di Maggio et al. 2019). In addition the regressions include dealer-day fixed effects in order to control for time variation in the tightness of balance sheet constraints of dealers – an important feature of this period (Duffie 2020).

F. Evidence from U.S. Treasuries

We additionally do some exploration regarding possible cross-country differences in the size-penalty. For example, should one expect the size penalty to be smaller in the U.S. Treasury market than in the U.K. gilt market? Given that the U.S. Treasury market is larger, deeper and more liquid than the U.K. gilt market (US-Treasury, 2021), inventory-based theories would be consistent with a smaller size penalty in U.S. Treasuries because of larger intermediation as well as interdealer activity (Viswanathan and Wang 2004). Similarly, recent information-based theories also predict that the size penalty would decrease in market size; see, for example, the model of Vives (2011) with strategic informed traders as well as the recent theory of market microstructure invariance (Kyle and Obizhaeva 2016).[46]

Comparing the U.S. and U.K. government bond markets is an ideal setting to estimate the effect of market size, as these two markets are similar in other aspects such as market structure and the riskiness of the assets. Measured in daily trading volume, the U.S. treasury market is around 10 times larger than the U.K. gilt market during our sample period.[47]

To estimate how the cost-size relation varies with market size, we use the ZEN data set on U.S. Treasuries which includes virtually all secondary market trades that are executed in London. Previous studies documented that at least around 3% of trading volume of the U.S. Treasury market occurs in London (Fleming 1997; Fleming, Mizrach, and Nguyen 2018). While the sample of Fleming (1997) for the London market is not representative of the whole Treasury market (e.g., trading volume in London is dominated by Treasury note trading and bill volume is extremely low overseas (p. 13 of Fleming 1997)), the recent samples in the ZEN and MIFID II data set are more representative (in terms of maturity) of the U.S. Treasury market (Ashtari et al. 2023).

Against this backdrop, we search for the same set of dealers and clients (as in our sample of U.K. gilts) and end up with around 1.3 million transactions for around 950 U.S. Treasuries. We find that about 75% of

[46] According to the theory of market microstructure invariance, for example, if the U.S. Treasury market volume is 10 times the U.K. gilt market volume and the volatility is the same in the two markets, then one would expect the largest one percent of U.K. gilt market trades to cost about $10^{1/3} \approx 2.15$ times as much as an equivalent trade in the U.S. Treasury market. We thank Pete Kyle for suggesting this example.

[47] During our baseline sample period (2011m8-2017m12), daily trading volume in U.S. Treasuries averages at around \$500 billion, whereas daily trading volume is around £30 billion in the U.K. gilt market (AFME 2018).

our sample of U.K. gilt market clients trade Treasuries in our sample of U.S. trades. Given that we use a similar set of dealers and clients in our empirical analysis, we also implicitly address the concern that the composition of clients and dealers may be endogenous to the market in question.

Table F.1
Trading Costs and Trade Size in Government Bond Markets: Evidence from U.S. Treasuries

	(1)	(2)	(3)	(4)	(5)
Trade Size	0.016	0.048*	0.047*	0.045*	0.050*
	(0.77)	(1.92)	(1.90)	(1.82)	(1.94)
N	1274295	1274289	1274289	1269855	1269238
R^2	0.055	0.061	0.062	0.134	0.139
Day#Bond FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer FE	No	No	Yes	No	No
Day#Dealer FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects (2), using our sample of U.S. Treasury transactions. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

We then re-estimate the five regression models as in our baseline Table 2 using our sample of U.S. Treasury trades. The results are presented in Table F.1. Column (1) shows that the cost-size relation is statistically insignificant without client fixed effects. Columns (2)-(5) show that adding client fixed effects generates a size penalty that is both statistically and economically weaker than what we find for U.K. gilts. Using the most conservative specification, we find that the size penalty in U.S. Treasuries is about one third of the size penalty in U.K. gilts (0.050 vs 0.158).[48]

[48] Figure F.1 in the Appendix shows that the cross-sectional relationship between trader size and trading costs is also weaker in U.S. Treasuries than in U.K. gilts. The size discount in U.S. Treasuries, measured as the slope coefficient, is around one fourth of the size discount in U.K. gilts (-0.16 vs -0.59 in Figure 3).

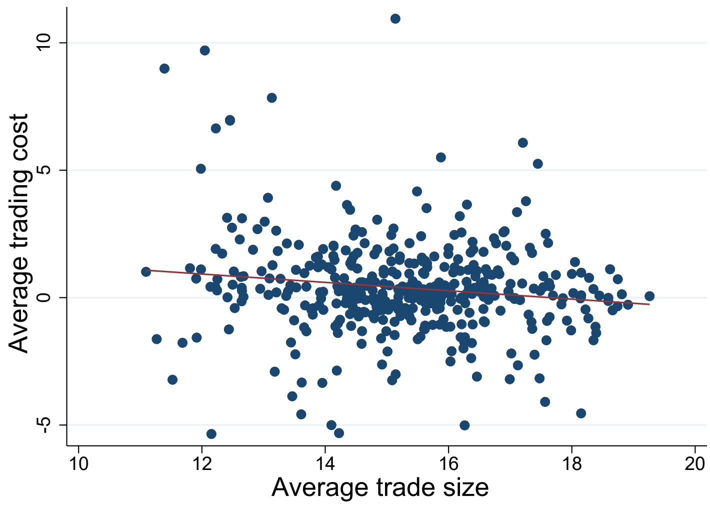


Figure F.1
Trading Costs and Trader Size in the Cross-Section: Evidence from the United States
This figure shows a scatter plot of average client trading costs (vertical axis) against average trade size (horizontal axis) at the client level in the U.S. government bond market. Average trading cost is the client-specific mean of our baseline cost measure 1. Average trade size is the natural logarithm of the average nominal size of a client's transactions. To reduce noise, the data set is trimmed at 1% level, leaving 586 observations. The estimated $\hat{\gamma} = -0.16$ with t-stat (based on robust standard errors) of -2.6.

Moreover, we also collapse our data at the client level to estimate the size discount in the cross-section of U.S. Treasury market clients. Figure F.1 estimates the scatter plot which gives a qualitatively similar picture to our baseline for the United Kingdom (Figure 3). However, the slope coefficient is four times as small for the U.S. sample (-0.16) as for the U.K. sample (-0.59), implying that the size discount is substantially stronger in the U.K. gilt market compared to our sample of U.S. Treasury trades.

G. Additional Tables

Table G.1
Summary Statistics on Trade Size

	(1) N	(2) Mean	(3) p10	(4) p50	(5) p90	(6) sd
			Trade Size (£s)			
Government Bonds						
All Clients	1274548	7825263	12856	850000	2.07e+07	4.56e+07
Less Sophisticated Clients	601157	8569644	15000	1000000	2.50e+07	4.58e+07
More Sophisticated Clients	673391	7160731	11000	600000	1.86e+07	4.55e+07
Corporate Bonds						
All Clients	1227954	1228126	9000	200000	2850000	6111960
Less Sophisticated Clients	561528	1283479	9000	100000	2600000	7350338
More Sophisticated Clients	666426	1181485	9000	263000	3000000	4827430

This table reports summary statistics for our baseline sample, covering the period from August 2011 to December 2017. Trade size is measured as the nominal size of the transaction in £s. The summary statistics is split based on client types (more sophisticated = asset managers + hedge funds; and less sophisticated = pension funds, insurance companies, foreign central banks, commercial banks, other non-financials) as well as markets (government bond vs corporate bonds).

Table G.2
Trading Costs and Trade Size in Government Bond Markets: Alternative Cost Measures

	(1)	(2)	(3)	(4)	(5)
Alternative Cost Measure I: Using Day-Bond-Dealer Average for $\overline{P}$					
Trade Size	-0.103***	0.061**	0.068***	0.077***	0.085***
	(-3.70)	(2.35)	(2.71)	(2.99)	(3.20)
N	973952	973948	973948	969689	968913
R^2	0.061	0.065	0.065	0.126	0.132
Alternative Cost Measure II: Using Day-Bond-Within Day Average for $\overline{P}$					
Trade Size	-0.180***	0.097***	0.115***	0.131***	0.149***
	(-3.80)	(3.75)	(5.14)	(6.04)	(7.33)
N	1261480	1261474	1261474	1256983	1256358
R^2	0.052	0.063	0.064	0.133	0.139
Alternative Cost Measure III: Using Day-Bond-Sell/Buy Average for $\overline{P}$					
Trade Size	-0.220***	0.100***	0.119***	0.132***	0.156***
	(-4.09)	(3.00)	(3.96)	(4.21)	(5.16)
N	1271266	1271260	1271260	1266824	1266209
R^2	0.003	0.010	0.011	0.087	0.093
Alternative Cost Measure IV: Using Average IDB prices for $\overline{P}$					
Trade Size	-0.184***	0.151***	0.173***	0.185***	0.212***
	(-3.41)	(4.47)	(5.58)	(5.91)	(6.91)
N	1232310	1232304	1232304	1227792	1227163
R^2	0.060	0.066	0.066	0.138	0.143
Day#Bond FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer FE	No	No	Yes	No	No
Day#Dealer FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client level. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

Table G.3
Trading Costs and Trade Size in Government Bond Markets Using Weighted Regressions: Alternative Cost Measures

	(1)	(2)	(3)	(4)	(5)
Baseline Cost Measure					
Trade Size	-0.321***	0.085**	0.119***	0.117***	0.146***
	(-9.27)	(2.26)	(3.30)	(3.90)	(4.76)
N	1274295	1274289	1274289	1269855	1269238
R^2	0.321	0.319	0.319	0.458	0.468
Alternative Cost Measure I: Using Day-Bond-Dealer Average for $\overline{P}$					
Trade Size	-0.127***	0.091***	0.102***	0.076***	0.090***
	(-5.47)	(2.97)	(3.34)	(2.70)	(3.10)
N	973952	973948	973948	969689	968913
R^2	0.313	0.303	0.303	0.443	0.456
Alternative Cost Measure II: Using Day-Bond-Within Day Average for $\overline{P}$					
Trade Size	-0.253***	0.065**	0.097***	0.110***	0.136***
	(-8.64)	(2.25)	(3.49)	(4.37)	(5.15)
N	1261480	1261474	1261474	1256983	1256358
R^2	0.319	0.321	0.322	0.459	0.469
Alternative Cost Measure III: Using Day-Bond-Sell/Buy Average for $\overline{P}$					
Trade Size	-0.324***	0.084**	0.118***	0.115***	0.144***
	(-9.35)	(2.23)	(3.27)	(3.83)	(4.72)
N	1271266	1271260	1271260	1266824	1266209
R^2	0.278	0.276	0.276	0.424	0.435
Alternative Cost Measure IV: Using Average IDB prices for $\overline{P}$					
Trade Size	-0.309***	0.122***	0.156***	0.153***	0.188***
	(-8.60)	(3.04)	(4.02)	(4.68)	(5.63)
N	1232310	1232304	1232304	1227792	1227163
R^2	0.317	0.319	0.320	0.457	0.467
Day#Bond FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer FE	No	No	Yes	No	No
Day#Dealer FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. Each observation is weighted by the inverse of the total number of transactions of the given client. The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

Table G.4
Trading Costs and Trade Size in Government Bond Markets: More Sophisticated Clients vs Less Sophisticated Clients, Using Alternative Cost Measures

	(1)	(2)	(3)	(4)	(5)
Alternative Cost Measure I: Using Day-Bond-Dealer Average for $\overline{P}$					
Trade Size#LessSophisticated	-0.132***	0.009	0.016	0.026	0.037
	(-3.98)	(0.21)	(0.37)	(0.63)	(0.86)
Trade Size#MoreSophisticated	-0.073*	0.108***	0.118***	0.125***	0.130***
	(-1.89)	(4.25)	(4.76)	(4.59)	(4.68)
p-values, eq. of coeff.	0.243	0.048	0.037	0.043	0.070
N	965894	965890	965890	958302	957486
R^2	0.109	0.112	0.112	0.195	0.201
Alternative Cost Measure II: Using Day-Bond-Within Day Average for $\overline{P}$					
Trade Size#LessSophisticated	-0.188***	0.052	0.064*	0.088**	0.111***
	(-4.10)	(1.35)	(1.75)	(2.55)	(3.10)
Trade Size#MoreSophisticated	-0.171**	0.130***	0.155***	0.168***	0.178***
	(-2.42)	(4.43)	(6.86)	(8.12)	(8.98)
p-values, eq. of coeff.	0.840	0.105	0.032	0.045	0.097
N	1257860	1257854	1257854	1251253	1250635
R^2	0.096	0.107	0.108	0.198	0.204
Alternative Cost Measure III: Using Day-Bond-Sell/Buy Average for $\overline{P}$					
Trade Size#LessSophisticated	-0.218***	0.057	0.067	0.079*	0.104**
	(-4.14)	(1.20)	(1.44)	(1.75)	(2.24)
Trade Size#MoreSophisticated	-0.218***	0.137***	0.166***	0.181***	0.195***
	(-2.73)	(3.29)	(4.81)	(5.09)	(5.65)
p-values, eq. of coeff.	0.997	0.203	0.085	0.072	0.115
N	1267885	1267879	1267879	1261353	1260735
R^2	0.051	0.058	0.059	0.160	0.165
Alternative Cost Measure IV: Using Average IDB prices for $\overline{P}$					
Trade Size#LessSophisticated	-0.191***	0.111**	0.120**	0.131***	0.159***
	(-3.57)	(2.15)	(2.39)	(2.60)	(3.01)
Trade Size#MoreSophisticated	-0.182**	0.175***	0.209***	0.226***	0.244***
	(-2.24)	(4.37)	(6.19)	(6.99)	(7.74)
p-values, eq. of coeff.	0.924	0.322	0.141	0.106	0.167
N	1230452	1230446	1230446	1223760	1223136
R^2	0.107	0.112	0.112	0.208	0.212
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer#ClientType FE	No	No	Yes	No	No
Day#Dealer#ClientType FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size interacted with client type dummies as well as various fixed effects (regression 4). The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.5
Trading Costs and Trade Size in Government Bond Markets: Big vs Small Macroeconomic News, Using Alternative Cost Measures

	(1)	(2)	(3)	(4)	(5)
Alternative Cost Measure I: Using Day-Bond-Dealer Average for $\overline{P}$					
Less Sophisticated Clients					
Trade Size#SmallNews	0.005	0.023	0.019	0.035	0.044
	(0.12)	(0.52)	(0.45)	(0.78)	(0.95)
Trade Size#LargeNews	0.018	0.024	0.010	0.015	0.028
	(0.39)	(0.54)	(0.23)	(0.33)	(0.60)
p-values, eq. of coeff.	0.610	0.960	0.723	0.476	0.565
More Sophisticated Clients					
Trade Size#SmallNews	0.114***	0.114***	0.110***	0.113***	0.122***
	(3.62)	(3.62)	(3.49)	(3.48)	(3.76)
Trade Size#LargeNews	0.127***	0.125***	0.139***	0.133***	0.141***
	(4.70)	(4.42)	(4.62)	(4.07)	(4.15)
p-values, eq. of coeff.	0.591	0.631	0.343	0.479	0.498
N	901677	897696	898657	894597	893772
R^2	0.112	0.142	0.164	0.194	0.199
Alternative Cost Measure II: Using Day-Bond-Within Day Average for $\overline{P}$					
Less Sophisticated Clients					
Trade Size#SmallNews	0.065	0.088**	0.072**	0.097***	0.120***
	(1.64)	(2.14)	(2.00)	(2.60)	(3.11)
Trade Size#LargeNews	0.065*	0.080**	0.065*	0.084**	0.110***
	(1.81)	(2.15)	(1.84)	(2.36)	(2.93)
p-values, eq. of coeff.	0.987	0.654	0.733	0.487	0.600
More Sophisticated Clients					
Trade Size#SmallNews	0.147***	0.164***	0.147***	0.161***	0.173***
	(5.51)	(6.46)	(5.70)	(6.81)	(7.64)
Trade Size#LargeNews	0.163***	0.173***	0.164***	0.170***	0.181***
	(6.12)	(6.52)	(6.35)	(6.91)	(7.53)
p-values, eq. of coeff.	0.422	0.627	0.386	0.633	0.649
N	1170316	1166823	1167791	1164241	1163626
R^2	0.107	0.138	0.169	0.196	0.201
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	No	No	No	No
Dealer#ClientType FE	Yes	Yes	Yes	No	No
Day#Dealer#ClientType FE	No	No	Yes	Yes	Yes
Month#Client FE	No	Yes	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size interacted with client type and macroeconomic surprise dummies as well as various fixed effects (regression 5). The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.6
Trading Costs and Trade Size in Government Bond Markets: Big vs Small Macroeconomic News, Using Alternative Cost Measures

	(1)	(2)	(3)	(4)	(5)
Alternative Cost Measure III: Using Day-Bond-Sell/Buy Average for $\overline{P}$					
Less Sophisticated Clients					
Trade Size#SmallNews	0.062	0.081	0.062	0.082*	0.104**
	(1.26)	(1.57)	(1.36)	(1.71)	(2.10)
Trade Size#LargeNews	0.067	0.068	0.064	0.073	0.105**
	(1.40)	(1.36)	(1.35)	(1.52)	(2.06)
p-values, eq. of coeff.	0.876	0.634	0.936	0.788	0.987
More Sophisticated Clients					
Trade Size#SmallNews	0.145***	0.156***	0.138***	0.151***	0.168***
	(3.91)	(3.96)	(3.44)	(3.57)	(4.08)
Trade Size#LargeNews	0.194***	0.206***	0.201***	0.205***	0.221***
	(4.56)	(4.81)	(4.80)	(4.84)	(5.23)
p-values, eq. of coeff.	0.112	0.099	0.082	0.130	0.134
N	1179331	1175860	1176855	1173330	1172712
R^2	0.058	0.090	0.129	0.157	0.162
Alternative Cost Measure IV: Using Average IDB prices for $\overline{P}$					
Less Sophisticated Clients					
Trade Size#SmallNews	0.125**	0.147**	0.126**	0.148***	0.175***
	(2.25)	(2.52)	(2.43)	(2.73)	(3.08)
Trade Size#LargeNews	0.118**	0.117**	0.113**	0.119**	0.153***
	(2.33)	(2.19)	(2.24)	(2.31)	(2.80)
p-values, eq. of coeff.	0.803	0.343	0.651	0.351	0.508
More Sophisticated Clients					
Trade Size#SmallNews	0.166***	0.182***	0.163***	0.177***	0.197***
	(4.16)	(4.33)	(4.10)	(4.31)	(4.86)
Trade Size#LargeNews	0.254***	0.267***	0.253***	0.260***	0.279***
	(5.95)	(6.10)	(6.37)	(6.45)	(7.00)
p-values, eq. of coeff.	0.017	0.017	0.033	0.049	0.052
N	1146316	1142804	1143736	1140154	1139534
R^2	0.112	0.141	0.179	0.205	0.210
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes	Yes
Client FE	Yes	No	No	No	No
Dealer#ClientType FE	Yes	Yes	Yes	No	No
Day#Dealer#ClientType FE	No	No	Yes	Yes	Yes
Month#Client FE	No	Yes	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size interacted with client type and macroeconomic surprise dummies as well as various fixed effects (regression 5). The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.7
Client Activity During Days with Big and Small Macroeconomic Surprises

	(1) Average Daily Transactions	(2) Average Daily Volume (£s)	(3) Average Daily Number of Clients	(4) Number of Days
Less Sophisticated Clients				
Small Surprise Days	361	3.46e+09	69	737
Big Surprise Days	391	3.83e+09	72	757
More Sophisticated Clients				
Small Surprise Days	402	3.24e+09	70	737
Big Surprise Days	433	3.49e+09	73	757

This table reports summary statistics on the activity of different client types on days with small and big macroeconomic surprises. The data covers the period from August 2011 to December 2017. The classification of small and big surprise days builds on the high-frequency methodology of Swanson and Williams (2014): we identify trading days when the surprise component of U.S. and U.K. macroeconomic announcements were high, by sort trading days into two groups, based on whether the magnitude of the surprise on day t was smaller or bigger than the sample median.

Table G.8
Trading Costs and Trade Size in Government vs Corporate Bond Markets: Using Alternative Cost Measures

	(1)	(2)	(3)	(4)
Alternative Cost Measure I: Using Day-Bond-Dealer Average for $\overline{P}$				
Less Sophisticated Clients				
Trade Size#GovernmentBonds	0.066*	0.079**	0.097**	0.108***
	(1.75)	(2.05)	(2.37)	(2.65)
Trade Size#CorporateBonds	0.201	0.112	0.107	0.113
	(1.50)	(0.76)	(0.71)	(0.72)
p-values, eq. of coeff.	0.247	0.797	0.940	0.970
More Sophisticated Clients				
Trade Size#GovernmentBonds	0.116***	0.127***	0.132***	0.134***
	(4.50)	(4.58)	(5.07)	(4.98)
Trade Size#CorporateBonds	0.377**	0.342**	0.360**	0.360**
	(2.52)	(2.22)	(2.23)	(2.17)
p-values, eq. of coeff.	0.074	0.142	0.142	0.155
N	790073	783440	783038	782305
R^2	0.360	0.433	0.437	0.440
Alternative Cost Measure II: Using Day-Bond-Within Day Average for $\overline{P}$				
Less Sophisticated Clients				
Trade Size#GovernmentBonds	0.099**	0.108***	0.141***	0.146***
	(2.28)	(2.71)	(3.70)	(3.83)
Trade Size#CorporateBonds	0.263	0.298*	0.299*	0.310*
	(1.53)	(1.72)	(1.70)	(1.70)
p-values, eq. of coeff.	0.268	0.216	0.318	0.316
More Sophisticated Clients				
Trade Size#GovernmentBonds	0.144***	0.169***	0.171***	0.177***
	(5.14)	(6.94)	(7.50)	(7.70)
Trade Size#CorporateBonds	0.730***	0.705***	0.717***	0.728***
	(4.56)	(4.67)	(4.62)	(4.55)
p-values, eq. of coeff.	0.000	0.000	0.000	0.000
N	1036375	1029952	1029616	1028996
R^2	0.357	0.436	0.441	0.445
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes
Client#Market FE	Yes	Yes	Yes	No
Dealer#Market#ClientType FE	Yes	Yes	Yes	No
Day#Dealer#ClientType FE	No	Yes	Yes	Yes
Month#Client FE	No	Yes	Yes	Yes
Client#Dealer	No	No	Yes	No
Client#Dealer#Market FE	No	No	No	Yes

This table regresses trading costs on trade size interacted with client type and bond market dummies as well as various fixed effects (regression 12). The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.9
Trading Costs and Trade Size in Government vs Corporate Bond Markets: Using Alternative Cost Measures

	(1)	(2)	(3)	(4)
Alternative Cost Measure III: Using Day-Bond-Sell/Buy Average for $\overline{P}$				
Less Sophisticated Clients				
Trade Size#GovernmentBonds	0.071	0.086*	0.124***	0.115**
	(1.33)	(1.86)	(2.79)	(2.50)
Trade Size#CorporateBonds	0.130	0.138	0.174	0.216
	(0.61)	(0.64)	(0.79)	(0.98)
p-values, eq. of coeff.	0.744	0.786	0.799	0.601
More Sophisticated Clients				
Trade Size#GovernmentBonds	0.147***	0.176***	0.187***	0.183***
	(3.81)	(4.68)	(5.19)	(5.08)
Trade Size#CorporateBonds	0.686**	0.674**	0.750***	0.854***
	(2.48)	(2.47)	(2.76)	(3.15)
p-values, eq. of coeff.	0.043	0.060	0.037	0.012
N	1054855	1048593	1048276	1047673
R^2	0.171	0.276	0.283	0.287
Alternative Cost Measure IV: Using Average IDB prices for $\overline{P}$				
Less Sophisticated Clients				
Trade Size#GovernmentBonds	0.111*	0.125**	0.167***	0.151***
	(1.74)	(2.43)	(3.29)	(2.88)
Trade Size#CorporateBonds	0.300	0.255	0.165	0.215
	(1.38)	(1.26)	(0.80)	(1.06)
p-values, eq. of coeff.	0.362	0.496	0.992	0.741
More Sophisticated Clients				
Trade Size#GovernmentBonds	0.193***	0.209***	0.230***	0.230***
	(5.11)	(5.36)	(6.06)	(6.10)
Trade Size#CorporateBonds	0.779***	0.767***	0.809***	0.876***
	(3.85)	(3.71)	(3.87)	(4.37)
p-values, eq. of coeff.	0.002	0.006	0.005	0.001
N	768755	761318	760936	760134
R^2	0.267	0.377	0.383	0.386
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes
Client#Market FE	Yes	Yes	Yes	No
Dealer#Market#ClientType FE	Yes	Yes	Yes	No
Day#Dealer#ClientType FE	No	Yes	Yes	Yes
Month#Client FE	No	Yes	Yes	Yes
Client#Dealer	No	No	Yes	No
Client#Dealer#Market FE	No	No	No	Yes

This table regresses trading costs on trade size interacted with client type and bond market dummies as well as various fixed effects (regression 12). The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.10
Trading Costs and Trade Size: Government vs Corporate Bonds Markets: All Clients Included

	(1)	(2)	(3)	(4)
Less Sophisticated Clients				
Trade Size#GovernmentBonds	0.077	0.083*	0.110**	0.116**
	(1.53)	(1.71)	(2.23)	(2.30)
Trade Size#CorporateBonds	0.447***	0.426***	0.353**	0.352**
	(2.87)	(2.90)	(2.43)	(2.40)
p-values, eq. of coeff.	0.007	0.008	0.055	0.067
More Sophisticated Clients				
Trade Size#GovernmentBonds	0.174***	0.197***	0.205***	0.206***
	(4.84)	(5.40)	(5.75)	(5.74)
Trade Size#CorporateBonds	0.670***	0.705***	0.743***	0.789***
	(4.03)	(4.48)	(4.72)	(5.03)
p-values, eq. of coeff.	0.002	0.000	0.000	0.000
N	1962998	1957464	1956891	1955799
R^2	0.287	0.350	0.356	0.358
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes
Client#Market FE	Yes	Yes	Yes	No
Dealer#Market#ClientType FE	Yes	Yes	Yes	No
Day#Dealer#ClientType FE	No	Yes	Yes	Yes
Month#Client FE	No	Yes	Yes	Yes
Client#Dealer	No	No	Yes	No
Client#Dealer#Market FE	No	No	No	Yes

This table regresses trading costs (measured in bp) on trade size (measured as the logarithm of the nominal size of the trade in £s) interacted with an indicator variable taking value 2 (1) if the trade takes place in the corporate (government) bond market. The regression also includes various fixed effects. The upper (lower) panel shows the results for less (more) sophisticated clients. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients, within a given client type.

Table G.11
Average Trade Size of More Sophisticated Clients Relative to Less Sophisticated Clients

	(1)	(2)	(3)	(4)
More Sophisticated Clients	-0.356	-0.162	-0.120	-0.085
	(-1.39)	(-0.76)	(-0.61)	(-0.58)
N	1274295	1274295	1273531	973952
R^2	0.149	0.282	0.350	0.566
Day#Bond FE	Yes	Yes	Yes	No
Dealer FE	No	Yes	No	No
Day#Dealer FE	No	No	Yes	No
Day#Bond#Dealer FE	No	No	No	Yes

This table regresses trade size on a dummy that takes the value of one if the client is more sophisticated (asset managers and hedge funds) or zero if they are less sophisticated (pension funds, central banks etc.) and various fixed effects. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

Table G.12
Trading Costs and Trade Size: Non-linearities

	(1)	(2)	(3)	(4)	(5)
Less Sophisticated Clients					
Trade Size $Q=2$	-0.061	-0.065	-0.021	-0.029	0.015
	(-0.51)	(-0.53)	(-0.20)	(-0.27)	(0.14)
Trade Size $Q=3$	0.247	0.221	0.224	0.225	0.285*
	(1.49)	(1.30)	(1.44)	(1.38)	(1.72)
Trade Size $Q=4$	0.353*	0.363*	0.356*	0.390*	0.479**
	(1.81)	(1.71)	(1.86)	(1.88)	(2.24)
N	598874	596569	597365	595007	594602
R^2	0.112	0.143	0.179	0.206	0.212
More Sophisticated Clients					
Trade Size $Q=2$	0.008	0.030	0.084	0.071	0.084
	(0.06)	(0.19)	(0.70)	(0.60)	(0.70)
Trade Size $Q=3$	0.242	0.257	0.301**	0.299**	0.319**
	(1.41)	(1.53)	(2.20)	(2.27)	(2.47)
Trade Size $Q=4$	0.664***	0.693***	0.705***	0.701***	0.744***
	(3.91)	(4.02)	(4.23)	(4.06)	(4.36)
N	672232	670823	670987	669573	669361
R^2	0.102	0.132	0.172	0.198	0.203
Day#Bond FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer FE	No	No	Yes	No	No
Day#Dealer FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The performance measures are in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

Table G.13
Trading Costs and Trade Size: Agency Trades of More Sophisticated Clients

	(2)	(3)	(4)	(5)
	Non-Agency Trades			
Trade Size	0.173***	0.194***	0.208***	0.222***
	(4.37)	(5.20)	(5.41)	(5.71)
	Agency Trades			
Trade Size	0.058	0.113*	0.078	0.079
	(0.88)	(1.72)	(1.26)	(1.29)
p-values, eq. of coeff.	0.074	0.257	0.049	0.027
N	656472	656472	647277	647029
R^2	0.159	0.160	0.282	0.286
Day#Bond	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	No	No
Dealer	No	Yes	No	No
Day#Dealer	No	No	Yes	Yes
Month#Client FE	No	No	Yes	Yes
Client#Dealer FE	No	No	No	Yes

This table regresses trading costs on trade size interacted with a dummy variable (taking value 1 if the given trade is an agency trade) as well as on various fixed effects. The performance measures are in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.14
Trading Costs and Trade Size: Client's Weak vs Strong (Top 1-2) Dealer Relationship

	(1)	(2)	(3)	(4)
Client's Weaker Dealer Relationships				
Trade Size	0.137***	0.164***	0.181***	0.211***
	(3.87)	(5.22)	(6.30)	(7.58)
Client's Strongest Dealer (top 1-2) Relationships				
Trade Size	0.069*	0.079**	0.093***	0.101***
	(1.87)	(2.41)	(2.97)	(3.19)
p-values, eq. of coeff.	0.026	0.000	0.001	0.000
N	1179684	1179684	1172979	1172350
R^2	0.110	0.110	0.204	0.209
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes
Client FE	Yes	Yes	No	No
Dealer#ClientType FE	No	Yes	No	No
Day#Dealer#ClientType FE	No	No	Yes	Yes
Month#Client FE	No	No	Yes	Yes
Client#Dealer FE	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The relationship measures are explained in Section 3.4. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p < .1$, ** $p < .05$, *** $p < .01$). The p-values correspond to the testing for the equality of coefficients.

Table G.15
Trading Costs and Trade Size: Dealers' Weak vs Strong (Top 1-2) Client Relationship

	(1)	(2)	(3)	(4)	(5)
Dealers' Weaker Client Relationships					
Trade Size	-0.196***	0.123***	0.137***	0.151***	0.176***
	(-4.32)	(3.51)	(4.42)	(4.70)	(5.64)
Dealers' Strongest (Top 1-2) Client Relationships					
Trade Size	-0.270**	-0.011	0.077	0.095	0.103
	(-2.01)	(-0.17)	(1.35)	(1.47)	(1.56)
p-values, eq. of coeff.	0.546	0.038	0.269	0.382	0.277
N	1170283	1170277	1170277	1160908	1160297
R^2	0.089	0.096	0.097	0.186	0.191
Day#Bond#ClientType FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer#ClientType FE	No	No	Yes	No	No
Day#Dealer#ClientType FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. The relationship measures are explained in Section 3.4. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$). The p-values correspond to the testing for the equality of coefficients.

Table G.16
Trading Costs and Trade Size in Government Bond Markets: Using Pre-Trade Benchmark Prices

	(1)	(2)	(3)	(4)	(5)
Trade Size	-0.213***	0.243***	0.269***	0.277***	0.294***
	(-3.08)	(4.61)	(5.30)	(5.09)	(5.27)
N	1199508	1199502	1199502	1194906	1194261
R^2	0.067	0.075	0.075	0.153	0.158
Day#Bond FE	Yes	Yes	Yes	Yes	Yes
Client FE	No	Yes	Yes	No	No
Dealer FE	No	No	Yes	No	No
Day#Dealer FE	No	No	No	Yes	Yes
Month#Client FE	No	No	No	Yes	Yes
Client#Dealer FE	No	No	No	No	Yes

This table regresses trading costs on trade size and various fixed effects. To measure transaction costs, we use a pre-trade benchmark price based on the lagged values of our second alternative benchmark price (footnote 14). Specifically, split the sample into three groups depending whether the transactions occur before 11 a.m., during 11 a.m.–3 p.m., or after 3 p.m. (thereby generating approximately even-sized subsamples in terms of number of transactions). We then compute (for each bond) the average transaction price in the given time window, and then use the benchmark price from the lagged window to compute transaction costs. The cost measure is in basis points. To reduce noise, we winsorize the sample at the 1% level. t-statistics in parentheses are based on robust standard errors, using two-way clustering at the day and client levels. Asterisks denote significance levels (* $p<.1$, ** $p<.05$, *** $p<.01$).

H. Additional Figures

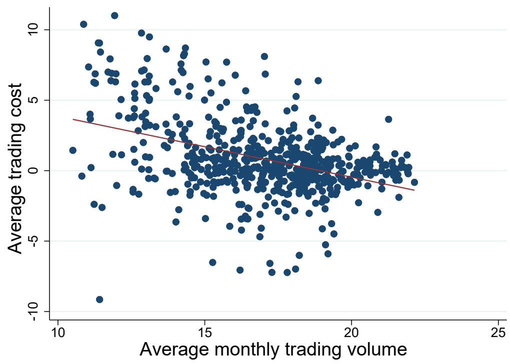


Figure H.1
Trading Costs and Trader Size in the Cross-Section
This figure shows a scatter plot of average client trading costs (vertical axis) against average trader size (horizontal axis) at the client level in the U.K. government bond market. Average trading cost is the unweighted mean of our baseline measure 1 at the client level. Trader size is measured as traders' monthly trading volume average across months. To reduce noise, the data set is trimmed at 1% level, leaving 586 observations. The estimated $\hat{\gamma} = -0.43$ with t-stat (based on robust standard errors) of -8.8.

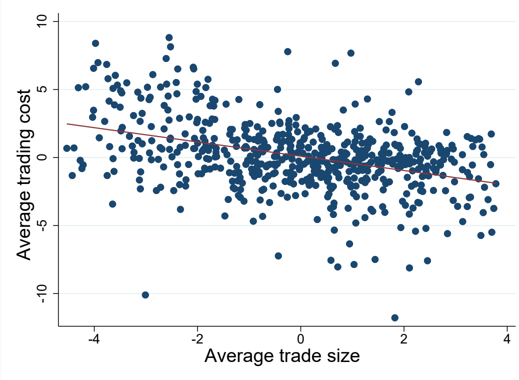
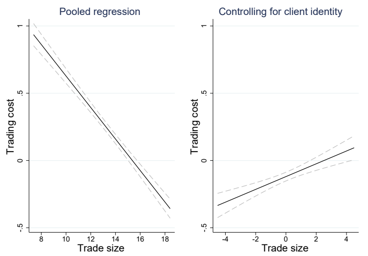
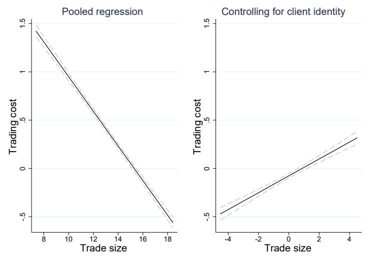


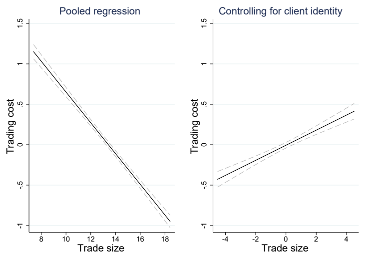
Figure H.2
Trading Costs and Trade Size in the Cross-Section: Adding Controls
This figure shows a scatter plot of average client trading costs (vertical axis) against average trade size (horizontal axis) at the client level in the U.K. government bond market. Average trading cost is the unweighted mean of our baseline cost measure 1 at the client level. Average trade size is the natural logarithm of the average nominal size of a client's transactions. To reduce noise, the data set is trimmed at 1% level, leaving 586 observations. The estimated $\hat{\gamma} = -0.52$ with t-stat (based on robust standard errors) of -9.2.



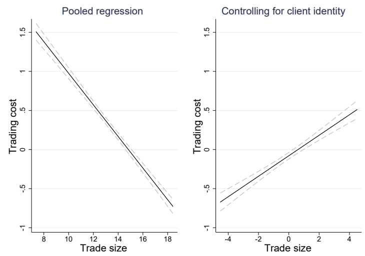
(a) Using Day-Bond-Dealer Average for $\overline{P}$



(b) Using Day-Bond-Within Day Average for $\overline{P}$



(c) Using Day-Bond-Sell/Buy Average for $\overline{P}$



(d) Using Average IDB prices for $\overline{P}$

Figure H.3
The Relation between Trade Size and Trading Costs: Using Alternative Cost Measures
The figures show a linear regression line on the pooled, transaction-level data (left panel) and on the data after we removed client-specific averages from trading costs and trade size corresponding to each trade. The four different trading cost measures are measured by 1 (building on O'Hara and Zhou (2021)) with different definitions of $\overline{P}$, and trade size is measured as the natural logarithm of the trade's notional. The four different performance measures are in basis points. The first measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t and dealer j. The second measure computes $\overline{P}$ in 1 as the average transaction price in bond k, in a given part of the trading day t. Using the time stamp for each trade, we divide trades into three groups, depending on whether the transaction occurred (1) before 11 a.m., (2) during 11 a.m.–3 p.m., or (3) after 3 p.m. The third measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, separately for buy and sell trades. The fourth measure computes $\overline{P}$ in 1 as the average transaction price in bond k, trading day t, using only trades on the interdealer market. The confidence bands are based on 95% standard errors as in Gallup (2019).