Essays in Market Microstructure

Universität Hamburg

Fakultät für Wirtschafts- und Sozialwissenschaften

Dissertation

Zur Erlangung der Würde des Doktors der Wirtschafts- und Sozialwissenschaften

(gemäß der Promotionsordnung vom 18. Januar 2017)

vorgelegt von

Neophytos Kathitziotis

aus Zypern

London, 2025

Erstbetreuer Prof. Dr. Ulrich Fritsche

Zweitbetreuer Prof. Dr. Dagfinn Rime

Datum der Disputation: 27.02.2026

Aknowledgements

This dissertation marks the end of a long journey. What began as an academic ambition gradually intertwined with the realities of professional and personal life. My early eagerness to explore a wide range of topics during my studies alongside the demands of a career in banking and the responsibilities of family life was often at odds with completing this work in a timely way. Yet the journey, with its detours and delays, brought lasting friendships, a deeper tacit knowledge of how financial markets operate, and the ability to approach research from different angles.

I would like to thank my supervisor, Ulrich Fritsche, for agreeing to supervise my dissertation, for giving me the freedom to shape my research independently, and for the patience to see it through at my own pace. Dagfinn Rime's willingness to step in as a reviewer of my dissertation, and his encouragement during the final stretch, have been immensely helpful.

My exposure to co-authors throughout this journey has equipped me with the skillset to thrive in academia, should I choose to pursue that path in the future. I am especially grateful to Carol Osler and Geir Bjønnes, who not only hosted me for research stays but also welcomed me into their families. Their willingness to engage deeply both professionally and personally helped me grow into the field of market microstructure and establish the core ideas of this thesis. Collaborating with Jo Saakvitne set a high standard for the remainder of my dissertation. Financial support from the Finance Markets Fund of the Norwegian Ministry of Finance facilitated these collaborations and part of the research underpinning this dissertation.

I am also grateful to a number of my former managers over the years for encouraging my professional development including funding training programmes and allowing me to take sabbaticals and internal research placements on the trading floor. These opportunities gave me the space to explore and deepen my understanding of market microstructure.

This thesis is empirical by design, a methodological choice. While methodology is not made explicit in the format of modern academic papers, I devoted significant time to studying economic methodology and attending related conferences, using these insights to inform my empirical approach. The topic deals with financial markets; the dataset is unique and exceptionally detailed and, where possible, I have tried to 'let the data speak freely'.

My wife, children, and family have endured this journey with me, sometimes with the thesis literally, and always figuratively in my backpack. I dedicate this work to my late mother and grandparents, who shaped the person I am today.

Table of Contents

List of figures **7**

List of tables **8**

Summary of thesis **10**

Zussamennfassung der Dissertation **15**

1 Chapter 1 – Introduction **20**

1.1 FX Market Microstructure and Trading Protocol 21

1.2 Liquidity Provision and Bid-Ask Spreads 22

1.2.1 Why do Adverse Selection Models Fail in OTC markets? 23

1.3 Information, Venue Choice and Price Discovery in FX markets 24

1.4 Conclusion 26

2 Chapter 2 - Cross-Client Variation in OTC Execution Costs **27**

2.1 Introduction 27

2.2 OTC Client Spreads: Literature Review 32

2.3 Data 33

2.3.1 Clients 34

2.3.2 Trading Platforms 35

2.4 OTC Markups: Estimating Models 36

2.4.1 Model 1: Perfect Competition and Dealer Costs 36

2.4.2 Model 2: Perfect Competition, Costs and Future Profits 37

2.4.3 Model 3: Market Power 38

2.5 OTC Markups: Model Estimates 43

2.5.1 Model 1: Costs under Perfect Competition 43

2.5.2 Model 2: Costs and Benefits Under Perfect Competition 44

2.5.3 Model 3: Market Power 45

2.5.4 Alternative Hypotheses 46

2.5.5 Robustness Tests 47

2.6 Cross-client Variation in OTC Liquidity Prices 49

2.6.1 Cross-client Variation in Liquidity Prices 49

2.6.2 Profit-seeking vs. Utilitarian Clients 50

2.6.3 Robustness 51

2.7 Concluding Remarks 51

3 Chapter 3 - Trading Venue Segmentation and Price Discovery in OTC Markets 59
3.1 Introduction 59
3.2 FX Trading landscape and venues....61
3.3 Information in FX markets 62
3.4 Data....64
3.5 Structural Vector Auto Regressions (SVARs) 67
3.6 Results 69
3.6.1 Cross Venue Interactions....70
3.6.2 How does price impact vary across client types....71
3.7 Conclusion....72
4 Chapter 4 - Frontrunning - Direct Evidence From a Global FX Liquidity Provider80
4.1 Introduction 80
4.1.1 Anticipatory trading....81
4.2 Model....84
4.2.1 Risk neutral dealer in a three-period model....87
4.2.2 The general model 88
4.3 Data....90
4.4 Methodology....92
4.4.1 Inventory analysis....92
4.4.2 Pre-trade price dynamics 94
4.5 Results 95
4.5.1 Pre-trade inventory effects....95
4.5.2 Pre-trade price effects....96
4.6 Robustness - are the timestamps reliable?....98
4.6.1 Time lag analysis - A method for validating time stamps in trade data 98
4.6.2 Results of time lag analysis 100
4.7 Descriptive statistics....102
5 Chapter 5 – Asymmetric information and otc bid-ask spreads....105
5.1 Data....109
5.1.1 Core-market Trades 109
5.1.2 Client Trades....110
5.1.3 Client Trading Platforms 111
5.2 Client Spreads and Core-Market Information Asymmetries....112
5.2.1 Regression Model....112
5.2.2 Main Results....113

5.3 Asymmetric Information Between Dealers and Clients: An Empirical Model......115
5.3.1 Client Information 115
5.3.2 Other Client Characteristics......116
5.3.3 Remaining Control Variables 119
5.4 Asymmetric Information Between Dealers and Clients: Results......121
5.4.1 Other Model Implications......123
5.4.2 Extensions and Robustness Tests 123
5.5 Information Chasing versus Adverse Selection 125
5.6 Concluding Remarks 126
6 Appendices......138
6.1 Appendix to Chapter 3......138
6.2 Appendix to Chapter 4......140
7 Bibliography 141
8 Declarations......150
8.1 Erklärung 150
8.2 Eidesstattliche Versicherung 150
8.3 Selbstdeklaration 151

List of figures

Figure 2.1.. Client Sophistication and Institutional Constraints 58

Figure 3.1: Equation 3.2 Impulse Response Functions - Venue Price Impacts........................77

Figure 3.2: Equation 3.2 Impulse Response Functions – Cross venue interactions 78

Figure 3.3: Equation 3.6 Client Price Impacts..79

Figure 4.1 Net Inventory Change around Large Trades 82

Figure 4.2: Timeline of trading activity..85

Figure 4.3: Timeline of the three-period model...87

Figure 4.4: Model-predicted inventory change for various degrees of risk aversion. 90

Figure 4.5: Customer spread versus market midprice..96

Figure 4.6: Estimated relation between trade size and pre-trade price impact 97

Figure 4.7 Theoretical RMSE Across Time Lags for Full Sample and Direct Trades 101

Figure 4.8 Empirical RMSE Across Time Lags for Full Sample and Direct Trades 102

Figure 5.1 EUR-USD Data sample realized volatility...136

Figure 5.*2* Client Sophistication and Institutional Constraints 137

LIST OF TABLES

Table 2.1:Descriptive Statistics.......53

Table 2.2: Determinants of FX Client Markups 54

Table 2.3: Determinants of FX Client Markups: Robustness Tests.......55

Table 2.4. Sources of Cross-client Variation in OTC Effective Spreads.......56

Table 2.5. Robustness Tests: Sources of Cross-client Variation in OTC Effective Spreads....57

Table 3.1: Descriptive statistics per client type 74

Table 3.2: Trade size in EUR millions by venue type 75

Table 3.3 Long-run impulse response.......76

Table 4.1: Mean pre-trade inventory changes and quantiles 95

Table 4.2: Pre-trade price impact regression models –.......98

Table 4.3: Pre-tradeprice impact regression models.......98

Table 4.4: Absolute volume by trade size.......102

Table 4.5: Number of trades by venue.......102

Table 4.6: Mean trade size by venue.......103

Table 4.7: Customer spread by trade size. 103

Table 4.8: Price change ($p_{t-30} - p_t$)) from t=-30 to 0 minutes by trade size. 103

Table 4.9: Customer spread by venue.......103

Table 4.10: Price change ($p_{t-30} - p_t$)) from t=-30 to 0 minutes by venue.......104

Table 5.1: Descriptive Statistics for Core-market Trades.......128

Table 5.2: Client Trades, Descriptive Statistics.......129

Table 5.3: Effective Core Spreads 130

Table 5.4: Client Information.......131

Table 5.5: Client Information and Client Markups.......132

Table 5.6. Robustness Tests.......133

Table 5.7. Trades with Negative Markups.......134

Table 5.8: Adverse Selection vs. Information Chasing 135

SUMMARY OF THESIS

"Cross-client Variation in OTC Execution Costs" with Geir Bjonnes and Carol Osler

This study provides a comprehensive analysis of variation in bid-ask spreads across OTC clients. We exploit the complete record of D2C trades by a major dealing bank in EUR-USD, a single contract for which daily volume exceeds daily volume in all US and European equities plus all US corporate and municipal bonds. These data have rare but important advantages for the study of OTC spreads: identifiers for individual clients and many client types, identifiers for over twenty trading platforms, and precise markups over the core-market (interbank) price.

EUR-USD spreads vary widely across clients. Spreads for the average SME or private wealth manager are an order of magnitude wider than spreads for the average hedge fund and the coefficient of variation for these OTC spreads is an order of magnitude larger than the coefficient of variation across S&P stocks (Hagströmer, 2021). Client spreads in FX forwards, which are likewise traded OTC, also vary widely (Hau et al., 2021).

The paper estimates three nested decomposition models tailored to liquid OTC markets. The model's dependent variable is the trade markup, which is superior to the effective spread for two reasons because it is the dealers' choice variable and because it is chosen after the dealer observes the other component of the client half-spread, the core-market half-spread. Model 1 assumes that spreads reflect dealer costs under perfect competition, where costs are higher for "direct" trades, which involve a human dealer, than for trades over electronic platforms. Model 2 assumes perfect competition but includes the possibility the dealers benefit from trading with active and informed clients (Bernhardt et al., 2005; Naik et al., 1999; Pinter et al., 2022). The third model assumes that dealers have market power relative to clients with limited negotiating skills or few potential counterparties (Duffie et al., 2005; Green et al., 2007).

The results indicate that dealer market power is the largest single contributor to cross-client variation in markups and thus spreads, accounting for at least 37%. The order-handling costs of direct trades are also important, contributing 16% outright. Market-power variables tend to be highest for clients that trade directly and their correlations jointly contribute a further 56%. These broad qualitative conclusions are sustained over numerous robustness tests and when clients are partitioned into profit-seeking and utilitarian clients, following Harris (2003).

Our analysis indicates that most cross-client variation would disappear if all clients fully exploited the available trading options, operated with maximum negotiating acumen, and maximized competition among dealers. Clients that pay high execution costs could be fully rational, nonetheless, achieving those standards is costly. It is costly to carry out the TCA required to reward employees for trading on low-cost platforms; it is costly to invest in trader negotiating skills; it is costly to establish and maintain new trading platforms; and it is costly to establish and maintain multiple dealing relationships. For the 40%-plus of FX clients in our sample that trade at most once per month, these costs could well exceed the associated savings.

Future research could productively examine the extent to which these findings apply in other OTC markets. Researchers could also fruitfully examine whether cross-client differences in OTC spreads have contributed to shifting financial market structures in recent decades.

"Trading Venue Segmentation and Price Discovery in OTC Markets"

This paper examines how trading venue segmentation shapes price discovery in the foreign exchange (FX) market. Drawing on a proprietary transaction-level dataset from a top-10 global dealer bank, the analysis spans more than 40 trading venues grouped into four key categories: interdealer order-driven, multi-dealer order-driven, request-for-quote (RFQ), and dark pools. Using structural vector autoregressions (SVARs), we are able to quantify the relative informational content of trades across venues and counterparty types. More specifically we offer a new perspective of how venue structure and participant behaviour may jointly influence price formation in OTC markets.

The study reveals that interdealer order-driven venues are the most informative, with trade-related price adjustments occurring within the first few minutes following execution. The estimated market impact from these venues is economically significant, representing approximately 40% of the representative bid-ask spread for an average interbank transaction. Trades on multi-dealer order-driven venues also contribute meaningfully to price discovery, albeit to a lesser degree. In contrast, our results suggest that bilateral RFQ venues serve predominantly as liquidity-providing mechanisms, order flow from these venues has a negative impact on midquote changes. This may reflect the nature of order flow that is less informed or directional, compared to activity on other trading venues. We also find that the main single-dealer platform acts as a risk warehousing venue, where inventory is initially accumulated rather than immediately offloaded to the market. This suggests that the dealer relies on internalization to manage short-term inventory imbalances before engaging external venues. One such external venue are Dark Pools, where trader anonymity ensures that price impacts will be small and short-lived.

When disaggregating order flow by counterparty type, we find that the most informative trades originate from high-frequency trading hedge funds and global liquidity providers. These entities operate primarily on transparent venues, and the corresponding SVAR estimates indicate that price adjustments occur within one to five minutes following their trades, reflecting the rapid incorporation of information into prices. In contrast, trades by customer banks and corporate clients are largely liquidity-driven and exhibit limited influence on price discovery.

This study advances the literature by offering the first comprehensive, venue-level assessment of FX price discovery that jointly considers venue segmentation and counterparty heterogeneity, two dimensions often analyzed in isolation in prior studies. Our findings have significant implications for market participants and policymakers, particularly regarding the effects of market fragmentation on price efficiency and trade execution costs.

"Order Anticipation and Large Trades in FX Markets" with Jo Saakvitne and Geir Bjonnes

This study provides novel empirical evidence on anticipatory trading by a major FX dealer bank in the period preceding large customer orders. Using a unique high-frequency dataset encompassing dealer-customer and interdealer transactions, we document systematic inventory adjustments by the bank in anticipation of large trades. The analysis reveals that, for orders exceeding €25 million executed via direct communication (e.g., phone, chat) or through the bank's proprietary platform (SBP), there is a significant build-up of dealer inventory up to 30 minutes before execution. This pattern is absent for smaller or more routine trades.

The study's findings suggest that dealers strategically position inventory ahead of large transactions, a behaviour often associated with "pre-hedging" or "frontrunning." The pre-trade adjustments are accompanied by distinctive price movements: market prices tend to move against the customer before execution, implying that anticipatory trading increases trading costs for large institutional clients. We frame these behaviours within a simple theoretical model, illustrating how profit-maximizing dealers engage in inventory management strategies that exploit predictable client order flow.

The empirical results are robust across different trade-size thresholds and highlight a key aspect of market microstructure in OTC FX markets: while anticipatory trading is restricted in regulated exchanges, its role in decentralized FX markets raises important questions regarding market efficiency, price discovery, and dealer conduct.

"Asymmetric Information and OTC Bid-ask Spreads" with Geir Bjonnes and Carol Osler

This paper examines the contributions of asymmetric information to the price of liquidity in quote-driven markets. We exploit highly detailed data from the world's largest OTC market, and make two contributions to the literature. First, the paper proposes and documents that quote-driven bid-ask spreads are influenced by two distinct information asymmetries: (i) the asymmetry between dealer and client and (ii) the asymmetry among professional traders in the core or interdealer market. This conclusion emerges from the two-tier nature of OTC markets and the fact that FX client prices are set as a markup over the same-side price in the core market (formerly interdealer market). The first information asymmetry influences the client markup, which the dealer sets relative to the same-side core-market price. The second information asymmetry influences the core-market half-spread.

A key contribution of this study is the empirical evidence that OTC dealers adjust bid-ask spreads in a more nuanced fashion than previously recognized. Rather than uniformly widening spreads to compensate for adverse selection risk, dealers selectively widen, narrow, or leave markups unchanged depending on the perceived information content of client trades.

In fact, we are able to show that dealers widen spreads for better-informed hedge funds, commodity trading firms, and client banks, consistent with adverse-selection risk (Glosten and Milgrom, 1985; Easley and O'Hara, 1987). Dealers narrow spreads for better-informed brokers, consistent with information chasing (Naik et al., 1999; Pintér et al., 2022). Finally, dealers make no adjustment to spreads for clients they consider uninformed, specifically low-leverage asset managers ("real-money funds"), MNCs, and SMEs.

We are able to uncover evidence for all three strategies because our comprehensive data include client identifiers and numerous client types and our empirical model of markup determination allows dealers to vary their strategy choices across client types. Further, that empirical model comprehensively includes the factors highlighted by the literature that are relevant to OTC markets.

The paper also examines how dealers adjust their strategic response to client information. Our evidence indicates that dealers correctly assess that real-money funds, MNCs, and SMEs are uninformed. With respect to the choice between adverse selection and information chasing, we hypothesize that HFT is a critical influence. Specifically, we argue that clients engaging in HFT may impose greater adverse-selection costs on dealers while providing dealers with fewer information-chasing benefits, relative to other clients. To test this hypothesis we partition the sample of hedge funds into those that do and do not engage in HFT and re-estimate the markup decomposition model. Results indicate that dealers widen spreads aggressively for better-informed HFT hedge funds and make little or no spread adjustments for non-HFT hedge funds, consistent with our hypothesis.

Future research could helpfully analyze the extent to which these conclusions hold in other OTC markets. It would be valuable to learn how dealers respond to client information in illiquid OTC markets without extensive HFT, such as those for US corporate or municipal bonds. It would also be valuable to learn whether interdealer information asymmetries influence interdealer spreads in illiquid OTC markets, and if so the extent to which those asymmetries influence client spreads.

ZUSSAMENNFASSUNG DER DISSERTATION

"Cross-client Variation in OTC Execution Costs" with Geir Bjonnes and Carol Osler

Diese Studie bietet eine umfassende Analyse der Variation von Geld-Brief-Spannen im Over-the-Counter-(OTC)-Handel über verschiedene Kundengruppen hinweg. Grundlage der Analyse ist der vollständige Handelsdatensatz einer führenden Dealerbank für EUR-USD-Geschäfte im Bereich „Dealer-to-Client" (D2C). EUR-USD stellt dabei ein Einzelkontrakt dar, dessen tägliches Handelsvolumen jenes aller US-amerikanischen und europäischen Aktien sowie sämtlicher US-Unternehmens- und Kommunalanleihen übersteigt.

Der verwendete Datensatz weist seltene, jedoch wesentliche Vorteile für die Analyse von OTC-Spreads auf: Er umfasst Identifikatoren für einzelne Kunden sowie Kundentypen, Angaben zu mehr als zwanzig Handelsplattformen sowie präzise Messgrößen des Preisaufschlags gegenüber dem jeweiligen Kernmarktpreis (Interbankenkurs).

Die Geld-Brief-Spannen im EUR-USD-Handel variieren erheblich zwischen verschiedenen Kundengruppen. Durchschnittliche Spannen für kleine und mittlere Unternehmen (KMU) sowie Vermögensverwalter im Privatkundengeschäft liegen um ein Vielfaches über jenen von Hedgefonds. Der Variationskoeffizient der Spreads im OTC-Bereich übersteigt jenen von S&P-Aktien um eine ganze Größenordnung (vgl. Hagströmer, 2021). Eine ähnliche Heterogenität lässt sich auch für Devisentermingeschäfte beobachten, die ebenfalls OTC gehandelt werden (vgl. Hau et al., 2021).

Die Studie schätzt drei ineinander verschachtelte Zerlegungsmodelle, die speziell auf liquide OTC-Märkte zugeschnitten sind. Die abhängige Variable in allen Modellen ist der Preisaufschlag je Transaktion, eine überlegene Maßgröße gegenüber dem effektiven Spread, da es sich hierbei um die unmittelbar vom Dealer bestimmte Preisentscheidung handelt, die zudem erst nach Kenntnis der jeweiligen Hälfte des Kernmarktspreads getroffen wird.

Modell 1 unterstellt perfekte Konkurrenz und interpretiert Spreads als Abbild dealerseitiger Transaktionskosten, die insbesondere bei „direkten" Handelsabschlüssen, also mit menschlicher Interaktion höher ausfallen als bei rein elektronischen Plattformen. Modell 2 bleibt ebenfalls im Rahmen perfekter Konkurrenz, berücksichtigt jedoch Informationsvorteile, die Händler durch Geschäfte mit aktiven oder informierten Kunden erzielen (vgl. Bernhardt et al., 2005; Naik et al., 1999; Pinter et al., 2022). Modell 3 geht darüber hinaus von asymmetrischer Marktmacht aus, die es Händlern erlaubt, gegenüber Kunden mit begrenzter Verhandlungsmacht oder eingeschränkter

Gegenparteiauswahl überdurchschnittliche Preisaufschläge zu realisieren (vgl. Duffie et al., 2005; Green et al., 2007).

Die Ergebnisse zeigen, dass die Marktmacht der Händler den größten Erklärungsbeitrag zur Variation der Preisaufschläge über Kundengruppen hinweg liefert mit einem Anteil von mindestens 37 %. Die Abwicklungskosten direkter Handelsgeschäfte tragen zusätzlich 16 % bei. Korrelationen zwischen Marktmachtmerkmalen und direkter Handelsabwicklung steigern den kombinierten Erklärungsbeitrag dieser Faktoren auf 56 %. Diese qualitativen Resultate sind robust gegenüber verschiedenen Modellvarianten und bleiben auch bestehen, wenn zwischen gewinnorientierten und nutzungsorientierten Kunden im Sinne von Harris (2003) differenziert wird.

Die Analyse legt nahe, dass ein Großteil der beobachteten Spreaddifferenzen verschwinden würde, wenn alle Kunden sämtliche verfügbaren Handelsoptionen vollständig ausschöpften, mit maximaler Verhandlungskompetenz agierten und konsequent Wettbewerb unter den Händlern erzeugten. Kunden mit hohen Ausführungskosten könnten dennoch rational handeln, da die zur Erreichung dieser Standards erforderlichen Investitionen beträchtlich sind: etwa für die Durchführung von „Transaction Cost Analysis“ (TCA), für Schulungen zur Verhandlungsführung, für die Etablierung und Pflege neuer Handelsplattformen sowie für den Aufbau paralleler Gegenparteibeziehungen. Für die über 40 % der analysierten FX-Kunden, die höchstens einmal pro Monat handeln, könnten diese Kosten die potenziellen Einsparungen sogar übersteigen.

Zukünftige Forschung könnte untersuchen, inwieweit sich diese Ergebnisse auf andere OTC-Märkte übertragen lassen. Ebenso wäre zu analysieren, ob Unterschiede in den Kundenaufschlägen zur strukturellen Veränderung von Finanzmarktarchitekturen in den letzten Jahrzehnten beigetragen haben.

"Trading Venue Segmentation and Price Discovery in OTC Markets"

Diese Arbeit untersucht, inwiefern die Segmentierung von Handelsplätzen die Preisentdeckung im Devisenmarkt (FX) beeinflusst. Grundlage der Analyse ist ein proprietärer Transaktionsdatensatz auf Einzelgeschäftsebene, der von einer globalen Top-10-Dealerbank bereitgestellt wurde. Der Datensatz umfasst mehr als 40 Handelsplätze, die in vier zentrale Kategorien gruppiert werden: interdealerbasierte Orderbuchmärkte, multiddealerbasierte Orderbuchmärkte, Request-for-Quote-(RFQ-)Plattformen sowie Dark Pools.

Mittels struktureller Vektorautoregressionsmodelle (SVAR) wird der relative Informationsgehalt von Transaktionen über unterschiedliche Handelsplätze und Kontrahententypen hinweg quantifiziert. Ziel ist es, ein neues Verständnis dafür zu entwickeln, wie die Struktur von

Handelsplätzen und das Verhalten der Marktteilnehmer gemeinsam die Preisbildung in OTC-Märkten determinieren.

Die Analyse zeigt, dass interdealerbasierte Orderbuchmärkte den höchsten Informationsgehalt aufweisen: Preisreaktionen erfolgen hier innerhalb weniger Minuten nach Ausführung der Transaktion. Der geschätzte Markteinfluss dieser Geschäfte ist ökonomisch signifikant und beträgt etwa 40 % der repräsentativen Geld-Brief-Spanne für ein durchschnittliches Interbankengeschäft. Auch Transaktionen auf multiddealerbasierten Orderbuchmärkten leisten einen substanziellen Beitrag zur Preisentdeckung, wenn auch in geringerem Umfang.

Demgegenüber deuten die Ergebnisse darauf hin, dass bilaterale RFQ-Plattformen primär der Liquiditätsbereitstellung dienen. Der von diesen Plattformen ausgehende Orderflow wirkt sich negativ auf Midquote-Veränderungen aus – ein Indiz dafür, dass dieser Fluss tendenziell weniger informationsbasiert oder richtungsweisend ist als Aktivitäten auf anderen Handelsplätzen.

Die Hauptplattform des Dealers fungiert darüber hinaus als Instrument zur Risikoübernahme: Hier werden Bestände zunächst intern absorbiert, bevor eine Weiterleitung an externe Märkte erfolgt. Dies spricht für ein gezieltes Inventory Management durch Internalisierung. Eine dieser externen Plattformen sind Dark Pools, deren anonyme Ausgestaltung dafür sorgt, dass Preiswirkungen gering und von kurzer Dauer sind.

Bei der Differenzierung des Orderflows nach Kontrahententyp wird deutlich, dass die informativsten Transaktionen von hochfrequenten Hedgefonds und globalen Liquiditätsanbietern ausgehen. Diese Akteure operieren primär auf transparenten Handelsplätzen. Die entsprechenden SVAR-Schätzungen belegen, dass Preisreaktionen innerhalb von ein bis fünf Minuten nach Ausführung eintreten ein Hinweis auf die schnelle Informationsverarbeitung durch den Markt.

Im Gegensatz dazu stehen Transaktionen von Kundenbanken und Firmenkunden, die primär liquiditätsgetrieben sind und kaum zur Preisentdeckung beitragen.

Die vorliegende Studie leistet einen neuartigen Beitrag zur Literatur, indem sie erstmals eine umfassende Analyse der Preisbildung im Devisenmarkt auf granularer Ebene des Handelsplatzes vornimmt unter gleichzeitiger Berücksichtigung von Segmentierung und Kontrahentenheterogenität, zwei Dimensionen, die in bisherigen Arbeiten meist isoliert betrachtet wurden. Die Ergebnisse sind von hoher Relevanz für Marktteilnehmer wie auch für Regulierungsbehörden, insbesondere im Hinblick auf die Auswirkungen von Marktfragmentierung auf Preiseffizienz und Ausführungskosten.

"Order Anticipation and Large Trades in FX Markets" with Jo Saakvitne and Geir Bjonnes

Die vorliegende Studie liefert neuartige empirische Evidenz zu antizipativem Handelsverhalten (*anticipatory trading*) einer führenden Dealerbank im Devisenmarkt im Vorfeld großvolumiger Kundenaufträge. Datengrundlage ist ein einzigartiger Hochfrequenzdatensatz, der sowohl Dealer-Kunden- als auch Interdealer-Transaktionen umfasst. Die Analyse dokumentiert systematische Bestandsanpassungen (*Inventory Management*) durch die Bank vor der Ausführung großer Transaktionen.

Für Aufträge über €25 Millionen, die über direkte Kommunikationskanäle (z. B. Telefon, Chat) oder über die unternehmenseigene *Single Bank Platform (SBP)* abgewickelt werden, zeigt sich ein signifikanter Aufbau von Händlerpositionen bis zu 30 Minuten vor Ausführung. Dieses Muster ist bei kleineren oder routinemäßigen Geschäften nicht zu beobachten.

Die Ergebnisse deuten darauf hin, dass sich Händler strategisch vor größeren Transaktionen positionieren – ein Verhalten, das in der Literatur häufig mit *pre-hedging* oder *frontrunning* assoziiert wird. Diese vorzeitigen Bestandsanpassungen gehen mit charakteristischen Preisbewegungen einher: Marktpreise tendieren vor Ausführung gegen den Kunden, was darauf hindeutet, dass antizipativer Handel die Handelskosten institutioneller Großkunden erhöht.

Diese Verhaltensmuster werden theoretisch im Rahmen eines Modells erläutert, das zeigt, wie gewinnmaximierende Händler vorhersehbare Kundenorderströme ausnutzen, um ihre Bestandsrisiken im Eigenhandel zu steuern.

Die empirischen Befunde erweisen sich als robust gegenüber unterschiedlichen Schwellenwerten der Handelsgrößen und beleuchten einen zentralen Aspekt der Marktstruktur im OTC-Devisenhandel: Während antizipatives Handeln an regulierten Börsen in der Regel verboten ist, wirft seine Rolle in dezentral organisierten FX-Märkten grundlegende Fragen zur Markteffizienz, zur Preisentdeckung und zum Verhalten von Händlern auf.

"Asymmetric Information and OTC Bid-ask Spreads" with Geir Bjonnes and Carol Osler

Diese Arbeit untersucht den Einfluss asymmetrischer Information auf den Preis von Liquidität in quotierten OTC-Märkten. Grundlage der Analyse ist ein äußerst detaillierter Datensatz aus dem größten OTC-Markt der Welt. Die Studie leistet zwei zentrale Beiträge zur bestehenden Literatur.

Erstens wird gezeigt, dass Geld-Brief-Spannen in quotengebundenen Märkten durch zwei unterschiedliche Informationsasymmetrien beeinflusst werden: (i) die Informationsasymmetrie zwischen Dealer und Kunde sowie (ii) die Informationsasymmetrie unter professionellen Händlern im Kernmarkt (früher: Interbankenmarkt). Dieses Ergebnis ergibt sich aus der zweistufigen Struktur von OTC-Märkten sowie der Tatsache, dass FX-Kundenpreise als Preisaufschlag auf den geld- bzw. briefseitigen Interbankenkurs festgelegt werden. Die erste Informationsasymmetrie beeinflusst den kundenbezogenen Preisaufschlag, den der Händler relativ zum Interbankenkurs derselben Seite festlegt. Die zweite Informationsasymmetrie beeinflusst die Hälfte des Spreads im Interbankenmarkt.

Ein zentraler Beitrag dieser Studie ist die empirische Evidenz, dass Dealer Geld-Brief-Spannen differenzierter anpassen als bisher angenommen. Anstatt Spreads pauschal zur Kompensation von adverser Selektion zu verbreitern, passen Händler die Preisaufschläge selektiv an, je nach wahrgenommenem Informationsgehalt der Kundengeschäfte.

Konkret zeigt sich, dass Händler Spreads gegenüber gut informierten Hedgefonds, Commodity-Trading-Unternehmen und Kundenbanken ausweiten, konsistent mit adverser Selektion (Glosten & Milgrom, 1985; Easley & O'Hara, 1987). Gegenüber gut informierten Brokern hingegen verengen Händler die Spreads, was mit *information chasing* im Sinne von Naik et al. (1999) und Pintér et al. (2022) übereinstimmt. Schließlich nehmen Händler gegenüber als uninformiert eingeschätzten Kunden – insbesondere institutionellen Anlegern ohne Leverage, multinationalen Konzernen (MNCs) und kleinen bis mittleren Unternehmen keine Spread-Anpassung vor.

Diese differenzierte Evidenz wird durch die außergewöhnliche Qualität der Daten ermöglicht: Der Datensatz enthält Kundenidentifikatoren sowie eine breite Typologie institutioneller Gegenparteien. Zudem erlaubt das zugrundeliegende Modell zur Analyse der Spreadkomponenten eine gezielte Schätzung händlerspezifischer Strategien nach Kundentyp.

Ein weiterer Schwerpunkt der Analyse liegt auf der strategischen Reaktion von Dealern auf wahrgenommene Kundeninformation. Die Ergebnisse belegen, dass Händler institutionelle Anleger, MNCs und KMUs zutreffend als uninformiert einschätzen. Bei der Unterscheidung zwischen adverser Selektion und *information chasing* identifizieren wir Hochfrequenzhandel (HFT) als maßgeblichen Einflussfaktor. Unsere Hypothese lautet, dass HFT-intensive Kunden bei Dealern höhere adverse Selektion verursachen, aber vergleichsweise geringe Vorteile in Bezug auf *information chasing* bieten. Zur Prüfung dieser Hypothese wird die Hedgefonds-Stichprobe in HFT-aktive und HFT-inaktive Gruppen aufgeteilt und das Modell zur Analyse der Spreadkomponenten erneut geschätzt.

INTRODUCTION

The last decade has witnessed the rapid adoption of information and communication technologies, particularly in financial markets and electronic trading. For instance, market estimates indicate that trading volumes in a wide range of financial products have risen between 10% and 40% (Conlin I.E.A 2023). The introduction of several regulations in Europe (MiFID II) and US (Dodd Frank) that amongst other goals require a shift from bilateral- to on exchange trading have further contributed to this trend.

Perhaps the most striking implication of these developments is the change of existing market structures and trading protocol arrangements (Bech et al. 2016). The once human-intermediated model of the floor specialist on the NYSE has now been replaced by a number of competing trading protocols ranging from all-to-all trading to dark pools and over the counter electronic trading. These changes have influenced the way market participants trade too. More investors can now access the market using a multitude of trading venues and in some instances compete with market makers for the provision of liquidity (Chaboud et al. 2023). The driving forces for these structural changes are complex and may differ across markets. Our focus is on the FX market, the largest financial market in the world (BIS 2022).

In order to understand why these issues are relevant let us examine the FX market's architecture. The FX market is highly decentralized and fragmented with trading occurring continuously on a large number of venues. In fact a recent report from BIS (2022) identifies at least 60 trading venues that dominate different segments of the market. When markets are decentralized without a common trading protocol, not all dealer quotes are observable to the market[1] and as a result market participants face uncertainty about the best available prices. In fact, in this dissertation we.study the customer transactions of a large FX dealer and find that bid-ask spreads differ significantly across client types, suggesting that market participants face heterogeneous trading costs depending on their trading characteristics and access to liquidity. Such costs may seem small given they are measured in increments of a cent. However, once the volume of such transactions in considered, their economic effect becomes non-trivial[2].

Recent changes in market structure have important implications for the behaviour of market participants too. Because electronic trading venues employ competing market clearing mechanisms,

[1] The level of observable pre-trade and post-trade information is more commonly referred to as market transparency
[2] For instance a 0,1 cent transaction cost in the 7.5 USD trillion (daily) FX market translates to 7.5bn USD of costs

market clearing will differ markedly across these venues: an (anonymous) order submitted to an exchange may have a very different price impact than a trade submitted to a RFQ platform (that is not anonymous); while some platforms guarantee trade execution some others don't and this uncertainty of execution will influence the speed at which new information is reflected in prices. Another important aspect of the 'electronification' of markets relates to how technology is rapidly impacting existing institutional arrangements and market structure. This trend is particularly pronounced in the FX market, where algorithmic trading has transformed execution practices and market access. According to BIS (2022), 75% of FX spot trading is algorithmic; therefore, technology, speed, and the ability to execute trades algorithmically has replaced the old dealer model of a handful of large banks. It is common now for trading venues to allow non-bank institutions to gain direct access to the interdealer market and in fact these providers are able to compete in liquidity provision with traditional dealers.

While such developments are important for academic discourse they have policy applications too. As regulation seeks to adapt to the technological transformation of trading, market microstructure has proven to be a useful tool in analyzing policy relevant outcomes. For instance the behaviour of dealers in anticipation of large transactions, known also as anticipatory trading is one topic that has attracted attention by regulators. While anticipatory trading is restricted in regulated exchanges—for example, through prohibitions on front-running under Regulation NMS in the U.S,.its role in decentralized FX markets raises important questions regarding market efficiency and dealer conduct. The microstructure agenda and this thesis in particular currently focus on these issues and namely the mechanisms that influence price formation, trading costs and more broadly the behavior of market participants. The following section introduces key features of the FX market's microstructure that motivate the empirical papers in this thesis.

1.1 FX Market Microstructure and Trading Protocol

The foreign exchange market is by far the largest OTC market with daily trade volume exceeding $7.5 trillion, with spot trading representing $2.1 trillion (BIS 2022). Financial participants have the lion share of spot trading, interdealer trading amounts to 52.5 percent and 45.5 percent is motivated by other financial participants, which includes hedge funds, investment managers, pension funds and public institutions. In contrast only 2 percent of the trades are intended to cover the transaction needs of firms importing or exporting products for use in production and trade. Trading takes place in two tiers, in one tier, dealers (banks) make markets by trading with clients; in the second tier, dealers trade with each other. Historically the dominant form of dealer trading has been bilateral through a voice broker, while nowadays it is largely anonymous through electronic venues as for example electronic brokers and dark pools. The majority of client trading is non-anonymous, namely clients need to reach out to a dealer before placing an order. Even when dealers provide a real-time stream of prices to clients on other venues, the trading process remains non-anonymous, the trading identities are disclosed before execution, creating a critical asymmetry in market functioning, because firstly this implies that dealers will know the identity of the client prior to trading and secondly, following a transaction, information about the client's trading intentions will remain private to the bank.

1.2 Liquidity Provision and Bid-Ask Spreads

In the previous section, we outlined some key characteristics of the market that are key to understanding how FX markets operate today. In this section, we briefly review the microstructure literature with an emphasis on the main models of bid-ask spread determination and provide some perspective of their relevance for OTC markets in particular.

Financial markets are designed to fulfil two principal functions: providing liquidity and enabling efficient price formation (O'Hara, 2003). The bid-ask spread is commonly understood as the price of liquidity. Microstructure research began by identifying the fundamental determinants of bid-ask spreads (Demsetz, 1968; Cornell, 1978; Copeland and Galai, 1983). It is commonly stated that spreads reflect three specific dealing costs: operating costs, inventory costs, and adverse selection. Strong supporting evidence for this traditional trinity came initially from studies of auction markets for equities (Glosten and Harris, 1988; Madhavan and Smidt, 1993; Huang and Stoll, 1997).

Order processing cost models claim that spread is the compensation for dealers who offer immediacy while bearing some fixed costs of market making. Many of the costs incurred by dealing rooms are fixed. These include the costs of providing the required physical infrastructure, the costs of providing information processing infrastructure including live news feeds and trading software; and the compensation of traders and related staff. Perhaps, the best known spread decomposition model is that of Roll (1984). In this model S measures the effective spread and Pt the transaction price at time t, which defines the effective spread, Sroll, as:

$$S_{Roll} = 2\sqrt{-\operatorname{cov}(\Delta P_t, \Delta P_{t-1})} \quad (1.1)$$

Equation (1.1) captures the pure bid-ask bounce, arising from order-processing costs and implies that the bid-ask bounce induces negative serial correlation in price changes. Hence, in a standard market setting with the presence of trading costs (bid - ask spreads), trades executed at the ask (bid) will create temporary upward (downward) price movements. Roll's model has another important implication, because a larger trade can cover a given fixed cost with a smaller proportionate spread, operating costs contribute a negative relation between spread and trade size.

The second class of market microstructure models study the activity of inventory management. Inventory risk models generally argue that the spread is the compensation for dealers who provide immediacy and assume risk by holding inventory at the same time. These models usually view dealers as risk-averse agents who provide liquidity and optimize their own security portfolios. The inventory costs literature identifies two key components: the cost of carry, which refers to the cost of holding inventory overnight, and the risk of capital gains or losses as asset prices fluctuate. According to the literature (e.g., Ho and Stoll, 1981), inventory risk rises with price volatility and with the size of a prospective transaction (because larger trades move the dealer further from desired inventory). Dealers manage the risk of inventory imbalance by adjusting their bid-ask quotes with the aim to induce trades in a favorable direction (Garman, 1976). This practice, known as quote shading, refers to the adjustment of bid or ask quotes to encourage directional trading and rebalance dealer inventory.

Asymmetric information or adverse selection models maintain that the bid-ask spread is the compensation for dealers who might lose money when trading with better-informed agents. Unlike inventory models, which emphasize the risks of holding positions, adverse selection models focus on the informational disadvantage dealers face when setting quotes in the presence of informed traders. According to Glosten and Milgrom (1985) rational dealers will widen spreads on the trades most likely to carry fundamental information. Easley and O'Hara (1987) show that informed traders have an incentive to trade larger amounts. In this case adverse selection concerns would rise with trade size, and if these concerns dominate the spread will widen with trade size. Information based models are supported by extensive evidence from the largest securities markets, namely bond and equity markets (e.g. Petersen and Sirri 2003).

1.2.1 Why do Adverse Selection Models Fail in OTC markets?

While the workhorse spread-decomposition models (Glosten and Harris, 1988; Madhavan and Smidt, 1993; Huang and Stoll, 1997) perform rather well in equity markets and more commonly anonymous market clearing mechanisms, they cannot diagnose the reasons why adverse selection consistently fails in OTC markets. One possible explanation is that bid-ask spreads are expected to vary between auction and OTC markets due to structural differences that alter the incentives and constraints of market participants. On most exchanges, liquidity providers move first, when they submit quotes for all market participants to see which means that the counterparty has not been determined when those quotes are disseminated. In OTC markets in contrast, the liquidity demander moves first by requesting quotes from a dealer. The difference in sequence creates a critical asymmetry in pre-trade anonymity: when quoting prices, market makers in auction markets do not know, and OTC dealers do know, the identity of their counterparty. Existing theory shows that OTC dealers can exploit this knowledge by price discriminating across at least three customer properties.

Firstly the customer's tendency to be informed about upcoming returns ("information") (Easley and O'Hara, 1987; Naik et al., 1999). According to this hypothesis, dealers will quote narrower spreads to potentially informed customers in order to learn from their information. Low transparency inherent in bilateral FX trading makes it possible for the bank to exploit the information in consecutive trades. These findings are also consistent with the literature on price discovery that claims that information in financial markets is incorporated into prices via the trading process (Menkhof and Schmeling 2010, Bjønnes et al 2021, Moore and Payne 2011, Osler et al 2011).

Secondly price discrimination across clients is also enabled by the instantaneous market power that arises from the fact that OTC markets are essentially search markets (Duffie et al. 2005) and customers see just one set of quotes, that is, one price of liquidity at a time and the process of searching for a better price is time-consuming and risky (Lagos and Rocheteau, 2009). Dealers can therefore extract higher mark-ups as customers weigh the costs of search against the spread on offer (Green et al., 2007).

Finally a third explanation relates to the customer's normal trading volume, because dealers will rationally use volume discounts to encourage future business from active customers (Bernhardt et al., 2005). Recent empirical results from FX markets (Hau et al. 2021) show that price discrimination might be better suited to explain such spread variation in OTC markets.

Together, these factors highlight that in OTC markets, dealers are not passive price takers subject to adverse selection, but active agents exploiting informational advantages and market power. This insight sets the stage for the next section, where we examine the role of information and dealer behaviour in price formation.

1.3 Information, Venue Choice and Price Discovery in FX markets

Information in the FX market is not symmetric, rather it is dispersed across a large number of participants and currencies (Ranaldo and Somogyi 2021). This dispersion means that no single participant holds complete information of the currency's value, and informed traders, such as hedge funds or institutional investors, will often possess private information that allows them to influence price formation. Consequently, price discovery in the FX market relies on the gradual aggregation of this fragmented information through trading activity. In this section we explore how theoretical models of information asymmetry and venue choice provide insights into these dynamics, setting the stage for the papers in this dissertation.

In the standard literature (Glosten and Milgrom 1985) information is revealed to the market only gradually through the trading process: uninformed participants will learn about the asset's true value by observing the trades of informed investors through time. The model assumes that when privately informed investors buy the asset, the price will increase, while the price will decrease when privately informed customers sell the asset.

However, compared with the "real-world" FX market, this model rests on several restrictive assumptions. First, it assumes the presence of a single dealer, whereas the FX market involves multiple dealers operating simultaneously. Second, the model posits that dealers themselves are uninformed, an assumption that conflicts with evidence suggesting that dealers often have access to private information they accumulate by observing customer order flow and through interdealer trading. These limitations impact the model's ability to fully capture the complexities and information dynamics of the FX market. The simultaneous trade-model by Evans and Lyons (2002) captures two important characteristics of the FX market that are overlooked in earlier models: (i) trading among dealers (interdealer trading), and (ii) private information among dealers. In round one, dealers accumulate private information by trading with clients and observing customer order flow. In round two dealers manage inventory and trade in the interdealer market. In round 3, dealers use information about trading in round 2 to set prices to the public. While this model captures well the interactions of dealers and customers in the FX market, it provides limited insights into the crucial aspect of venue choice. Understanding why participants select specific trading venues, whether due to trade size or information asymmetry is critical for understanding price discovery.

Henderschott and Madhavan (2015) study corporate bond markets that share a similar market structure to FX. They provide a framework that attempts to narrow the gap between theoretical models presented earlier and real market dynamics where varying trading protocols co-exist. They study venue choice among traders, in auctions (electronic) dealers compete to provide bids and there is pre-trade and post-trade transparency. In search markets (OTC), dealers interact with traders bilaterally. A key insight is that traders will prefer auction markets for smaller trades where immediate execution will be preferred while search markets are more useful for larger transactions where traders would like to keep information private and limit market impact. By shedding light on the relevance of trading protocols, these findings provide a framework for analyzing how fragmented markets like foreign exchange balance transparency, liquidity provision, and market impact.

1.4 Conclusion

This dissertation contributes to the growing literature on market microstructure by examining how trading protocols, information asymmetries, and dealer behaviour shape liquidity provision and price discovery in the FX market. Drawing on a unique dataset, the four papers that follow explore heterogeneity in bid-ask spreads across client types, the role of trading protocols and venue choice in price discovery, anticipatory trading by dealers, and the contributions of information asymmetries on the price of liquidity. Together, these studies provide a comprehensive view of how technological and institutional changes are reshaping the FX market, offering insights relevant not only to academics but also to policymakers and market practitioners.

Cross-Client Variation in OTC Execution Costs

2.1 Introduction

This paper analyzes how bid-ask spreads vary across clients in over-the-counter (OTC) markets. The spread, or equivalently the price of liquidity, is a crucial determinant of market efficiency and price discovery, and many studies have examined how spreads vary across stocks or bonds.[3] Few studies, however, have examined how spreads vary across traders. This reflects in part the early availability of data on order-driven markets, where all traders face the same spread at a given moment. In OTC markets, by contrast, liquidity suppliers rationally vary bid-ask spreads across clients.[4] Datasets from OTC markets are now readily available but few include client identifiers, so studies of cross-client spread variation remain rare (Hendershott et al., 2020; Hau et al., 2021).

Our analysis exploits the complete record of dealer-client transactions at a top-10 dealing bank in spot EUR-USD. In addition to shedding light on OTC markets generally, the analysis sheds light on a contract that merits study on its own. Daily trading in spot EUR-USD was roughly $616 billion in 2022, more than total daily trading in all US equities plus all European equities, plus all US corporate bonds, plus all US municipal bonds.[5] Our transactions data include price, quantity, and time, as is standard, plus client identifiers, many client types, identifiers for a plethora of trading platform, information on trades via prime brokers, and each trade's precise markup or markdown relative to the same-side "prevailing market price" (MSRB, 2018; (henceforth, markup). We consider seven client types: hedge funds; client banks; brokers; low-leverage asset managers such as mutual funds, pension funds, and endowments; managers of private wealth; MNCs; and SMEs. The detail of our data, which exceeds other studies of FX (e.g., Menkhoff et al., 2016; Ranaldo and Somogyi, 2021), proves indispensable for identification.

EUR-USD spreads vary dramatically across clients. Averages range from 0.7 bp for hedge funds to 38.7 bps for private clients. The coefficient of variation across clients, 1.02, exceeds by an

[3] Glosten and Harris (1988); Huang and Stoll (1997); Harris and Piwowar (2006); Edwards et al., (2007).
[4] Naik et al. (1999); Duffie et al. (2005); Bernhardt et al. (2005); Green et al. (2007); Pintér et al. (2022).
5 EUR-USE: Bank for International Settlements (2022). US equities: five-day mean of $161 billion, 4/3/2023. CBOE Global Markets (2023a). European equities: five-day mean of €41 billion, 4/3/2023. CBOE Global Markets (2023b).

US corporate bonds: five-day mean, $47 billion, 4/6/2023. Securities Industry and Financial Markets Association (2023a). US municipal bonds: five-day mean, $13 billion, 4/6/2023. Securities Industry and Financial Markets Association (2023b).

order of magnitude the coefficient of variation across S&P stocks, 0.7 (Hagströmer, 2021). OTC client spreads also vary widely in FX derivatives, where mean spreads for clients at the 25th and 75th percentiles are 2.5 pips and 30.1 pips, respectively (Hau et al., 2021).[6]

Our analysis shows that variation in OTC client spreads is not strongly influenced by inventory costs and adverse selection costs, though these are important for variation in spreads across stocks (Glosten and Harris, 1988; Huang and Stoll, 1997). Nor is it strongly influenced by two client characteristics known to elicit first-order price discrimination: trading volume (Bernhardt et al., 2005) and private information (Naik et al., 1999; Pintér et al., 2022).

Over 90% of the variation in OTC spreads would disappear if all clients consistently traded with maximum negotiating acumen and strategic flexibility. That is, if all clients consistently chose the least-costly trading platforms, negotiated skillfully, and declined to trade when quoted a wide spread, forcing the dealer to compete for their business. But this does not describe the trading of many small firms, non-financial firms, or firms that rarely trade FX. Employees who trade for these institutions often request quotes using email, and txt, an approach known as "direct trading" that is more costly than trading over fully-electronic platforms because it requires the time of a human dealer. Employees who trade for these institutions may not recognize a wide spread or know how to negotiate a better one, which gives dealers market power. Dealers also have market power relative to clients with few or no alternative counterparties. These clients are relatively insensitive to the price of liquidity and they will pay relatively high spreads in equilibrium (Ho and Stoll, 1981) even when multiple dealers compete (Duffie et al., 2005).

We provide evidence that differences in clients' reliance on direct trading, negotiating skills, and access to alternative counterparties mirror differences in their incentives to minimize execution costs. Negotiating acumen is achieved via a substantial investment in knowledge, skill, and technology, but over one-third of clients in our data trade less than once per week, so FX execution costs can be too small to warrant that investment. At non-financial clients that trade frequently, the employees that trade generally face no incentive to minimize execution costs so many choose trading platforms based on convenience. Likewise, achieving an additional FX dealing relationship involves the investment of time and effort required to qualify for a new line of credit. These investments need not be cost-effective for all clients.

These conclusions emerge from a set of nested empirical models tailored to the OTC setting. The first two assume perfect competition; the third assumes that dealers have market power (Duffie et al., 2005). A critical OTC-specific dimension of our models is the dependent variable, the client-

[6] A pip or "price improvement point" is one unit of an exchange rate's fifth significant digit. For EUR-USD a pip is $0.0001/€ or 0.8 basis point (bp) at the sample-average exchange rate of €1.31/$.

specific markup. Given the two-tier structure of OTC trading, each client half-spread comprises this markup plus the core-market (or interdealer) half-spread. It is econometrically appropriate to analyze markups separately because it they are set after dealers observe the core half-spread and in a different competitive setting.[7] A primary focus on markups is appropriate For our study, in particular, because markups generate over 99% of cross-client spread variation. In most existing studies of OTC spreads the dependent variable is the spread itself, measured either directly (Hendershott et al., 2020; Collin-Dufresne et al., 2020; Hau et al., 2021) or as returns across sequential trades (Harris and Piwowar, 2006; Edwards et al., 2007; Green et al., 2007; Goldstein et al., 2007; Feldhütter and Poulsen, 2018).

Our first model includes three major dealing costs that drive cross-stock variation in order-driven markets – order-handling costs, inventory costs, and adverse selection costs (e.g., Glosten and Harris, 1988; Huang and Stoll, 1997; Harris and Piwowar, 2006) – all of which can vary across OTC clients. Order-handling costs depend on the client's choice of trading platforms and are higher for direct trades. Inventory costs can depend on the client's trade sizes (Stoll, 1978) and trade timing (Tkatch and Kandel, 2008). Adverse-selection costs can vary with an OTC client's information, because dealers know their potential counterparty when quoting: each OTC trade begins when a client requests quotes from a specific dealer, thereby eliminating anonymity. OTC dealers become familiar with each client, including whether they are informed, during long-term dealing relationships.

Estimates of the cost-only model, estimated using OLS with Newey-West standard errors clustered by date, confirm that order-handling costs are substantially higher when clients trade directly. Otherwise, however, the results are inconsistent with standard theories. Contrary to inventory theories, markups are inversely related to volatility (Stoll, 1978) and inversely related to the waiting time between trades (Tkatch and Kandel, 2008). Contrary to adverse-selection theory, client information is insignificant (e.g., Easley and O'Hara, 1987).

Model 2 continues to assume perfect competition but permits dealers to use narrow spreads to attract clients whose future business would be profitable. Dealers evidently profit from the future business of active clients (Bernhardt et al., 2005). Dealers can also profit from learning an informed client's trade direction, information they can use to adjust quotes to others (Pinter et al., 2021) or to take positions before a client's information becomes embedded in the price (Naik et al., 1999). Evidence for this "information chasing" strategy in FX is provided in Ramadorai (2008). Unlike that study, however, Model 2 allows dealers to vary their strategic response to information – information chasing, adverse selection, no response – across client types.

[7] Core markets for liquid FX currency pairs are largely order-driven.

Estimates from Model 2 support the hypothesis that dealers seek to attract clients with profitable characteristics. As predicted in Bernhardt et al. (2005), dealers quote narrower spreads for relatively active clients. As predicted in Pintér et al. (2022) and Naik et al. (1999), dealers quote narrower spreads to better-informed brokers. The flexibility of Model 2 enhances our insights into OTC dealer responses to client information in two ways. First, it provides positive evidence for adverse selection: dealers quote wider spreads for better-informed hedge funds and client banks. Second, Model 2 provides evidence that dealers make no adjustment for information asymmetries for private clients, real-money funds, and non-financial clients. This approach is logical given evidence here and elsewhere that these clients tend to be uninformed (Menkhoff et al., 2016; Ranaldo and Somogyi, 2021; Bjønnes et al., 2021).

Model 3 assumes that dealers have market power with respect to clients that lack negotiating skills or alternative counterparties (Duffie et al., 2005). OTC prices are negotiated: a dissatisfied client can express concern or, more commonly, decline to trade and seek better quotes with another dealer. Clients with limited sophistication or few alternative counterparties will be quoted wider spreads because they have little bargaining power (Duffie et al., 2005; Green et al., 2007; Hendershott et al., 2020; Hau et al. 2021). Further, these clients can pay wider markups even when quoted a normal spread because dealers skew the quotes when trade direction is predictable – e.g., raising both quotes for likely EUR purchase – expecting the distortion to go unrecognized.

Dealers with market power can also adopt a strategy we introduce here, intertemporal smoothing. FX dealers know that their clients – like customers in any market – find price changes to be "confusing, frustrating, and annoying" (Dholakia, 2016, p. 1). Dealers are also aware that quoting a consistent price for liquidity in volatile markets is a sign of reliability (Cheung and Chinn, 2001). We hypothesize that rational OTC dealers encourage client loyalty by smoothing spreads over time within each client relationship.

Estimates of Model 3 provide extensive confirmation that OTC dealers achieve higher markups when they have market power. Markups are inversely related to client sophistication, and especially so for clients whose trade direction is predictable. Markups are higher for clients with few available counterparties. Evidence consistent with intertemporal smoothing comes from the markups' inverse relation to volatility and to the core spread's intraday pattern.

Our analysis of OTC client spreads finishes by quantifying the relative contributions of costs, benefits, and market power to cross-client markup variance. Client trading activity, client information, and intertemporal smoothing – all highly significant in the regressions – jointly contribute less than 3% of such variation. Two market-power variables – sophistication and potential counterparties – prove to be the biggest drivers of cross-client markup variance, contributing 37% in total, including their covariance. Also important are order-handling costs, the variance of which accounts for 16% of cross-client markup variance. A further 39% of cross-client markup variance is arises from strong cross-client correlations between order-handling costs and market-power variables: expensive direct trading is preferred among clients with the least incentive to minimize execution costs and thus the least negotiating leverage vis-à-vis their dealers. Our qualitative results are sustained when the model is re-estimated with alternative measures of four key independent variables, when EUR buys can be priced differently than EUR sales, and for a small subset of trades priced by human interbank dealers.

Our OTC spread decomposition model is flexible and could readily incorporate additional client characteristics. Nonetheless, the characteristics already included appear to span those highlighted in the literature. Client trading activity with our bank should encompass the strength of the dealer-client (D2C) trading relationship (Hau et al., 2021). Client credit risk is controlled in spot FX with credit limits, prime brokerage contracts, and margins instead of spread adjustments.

The wide cross-client variation in OTC spreads implies disproportionality in the burden of supporting market infrastructure. SMEs and private clients in our data account for 1.0% of traded value but 19.4% of bank markup income; client banks, by contrast, account for 51.5% of traded value but 29.7% of markup income. This disproportionate burden can help explain a shift among small clients from the major dealing banks to retail forex brokers beginning around 2000. This shift contributed to a sharp contemporaneous decline in the number of FX dealers.

More generally, our analysis can help explain instability in financial-market structures. Because OTC spreads vary widely, a substantial fraction of clients in any market could prefer a different market structure. In an OTC market, those who pay high spreads would prefer order-driven trading, where everyone pays the same spread. In order-driven markets, those who would benefit from lower OTC spreads would prefer OTC trading. To illustrate the latter, roughly a century ago almost all trading in US corporate and municipal bonds took place on the NYSE (Biais and Green, 2007). After 1929, however, US corporate bond trading shifted from the exchange to OTC. Biais and Green (2007) suggest that this shift, which was costly to retail traders, was driven by institutional traders who could expect narrow spreads in OTC trading.

The rest of this paper has five sections and a conclusion. Section 2.2 reviews relevant literature. Section 2.3 describes our data. Section 2.4 develops the three markup decomposition models. Section 2.5 presents the estimation results. Section 2.6 uses the estimated coefficients to quantify the sources of cross-client spread variation. Section 2.7 concludes.

2.2 OTC Client Spreads: Literature Review

Though there are many theoretical and empirical studies of OTC spreads,[8] few examine the potential for variation across clients. That possibility has often been suggested, nonetheless, as an explanation for the consistently negative relation between trade size and OTC spreads.[9] This relation could, in theory, reflect nothing more than fixed order-handling costs (Green et al., 2007; Edwards et al., 2007). It has been considered puzzling, nonetheless, because it seems to violate adverse selection, for which abundant evidence has emerged in equity markets (e.g., Glosten and Harris, 1988; Huang and Stoll, 1997). Easley and O'Hara (1987) show that if dealers can adjust the spread for each trade, as in OTC markets, larger trades should pay wider spreads because they are more likely to be informed. The FX market provides further evidence that seems to violate adverse selection: financial clients, who are better informed than non-financial clients, pay narrower spreads (Osler et al., 2011).

Theoretical explanations for these findings have pointed to the ability of OTC dealers to vary spreads strategically in response to client characteristics that could be correlated with trade size or client type: trading activity, private information, and sophistication. OTC dealers could rationally set narrower spreads for active clients, to attract their future business (Bernhardt et al., 2005). OTC dealers could rationally quote narrower spreads for better-informed clients whose information they can exploit via parallel interdealer trades (Naik et al., 1999) or pro-active adjustments to quotes for others (Pintér et al., 2022). Finally, dealers could quote narrower spreads to clients whose low search costs (Duffie et al., 2005) or "sophistication" (Green et al., 2007) limits the dealers' market power. The puzzling negative relation between trade size and spreads could emerge if more active, better informed, or more sophisticated clients tend to make the larger trades.

Unfortunately, if clients engage in stealth trading (Barclay and Warner, 1993) trade size may not be monotonically related to any of these characteristics. Splitting large trades is standard practice in FX, and in the U.S. corporate bond market many large and sophisticated institutions make many small trades, according to evidence in O'Hara et al. (2018). Further, omitted-variable bias could distort coefficients if trade size proxies for client characteristics.

Testing the empirical relevance of client characteristics for OTC spreads has been impeded by the absence of client identifiers in OTC datasets. Even studies that benefit from client identifiers can find it difficult to identify all of the relevant characteristics. O'Hara et al. (2018) use data with

---

[8] Bessembinder et al. (2006); Goldstein et al. (2007); Harris and Piwowar (2006); Edwards et al. (2007); Collin-Dufresne et al. (2019); Collin-Dufresne et al. (2020).

[9] These include the London Stock Exchange (Reiss and Werner, 1996); the modern U.S. corporate bond market (Goldstein et al., 2007; Edwards et al., 2007; Hollifield et al., 2017); the modern U.S. municipal bond market (Harris and Piwowar, 2006; Green et al., 2007), the pre-war corporate bond market (Biais and Green, 2007), and the foreign exchange market (Osler et al., 2011)

client identifiers to provide evidence of activity discounts in U.S. municipal bond markets. However, they do not control for client sophistication or information. Hendershott et al. (2020) and Hau et al. (2021) use data with client identifiers to provide evidence that client activity and their number of FX dealing relationships influences OTC spreads, but they do not control for client information or sophistication.[10] Our analysis accounts explicitly for all four of the client characteristics highlighted in the literature.

Data limitations have also led most existing studies to use effective spreads, rather than markups, as dependent variable. Because dealers respond to the core spread when setting the markup this can confound the identification of dealer strategies.

2.3 Data

Our data come from the world's largest OTC market, FX, where daily trading exceeds \$7.5 trillion (Bank for International Settlements, 2022). The primary dataset comprises the complete front-office transaction record of spot EUR-USD trades by a top-10 bank over 68 trading days from 2 January to 20 April 2012. Secondary data are bid and ask quotes sampled at the one-second frequency from Reuters' order-driven core-market platform.[11]

The bank's trade records include basic descriptors – quantity, price, time – plus trade initiator, trade direction, precise markup, client identifiers, client type assigned by the bank, trading desk, client country of origin, trading platform, a flag for trades priced by an interbank dealer, and details of trades on which the bank serves as prime broker. Trade time is reliable and measured to the second. Information on all trades is complete. As at most large dealing banks, the majority of client trades are internalized.

The bank calculates the markup as follows: $Markup_t = D_t\ (p_t^{client}\text{-}p^{core})$, where D_t=1 (-1) for client purchases (sales) and p^{core} is the bank's adjusted version of the contemporaneous same-side core price. Adjustment is required because many core platforms permit liquidity providers to decline to deal after a client accepts posted quotes (an option known as "last look"), but the bank's client quotes must be firm. The bank identifies firm-equivalent core prices by adjusting in real time for the state of the market and the possibility of execution failure on core platforms.

[10] Hau et al. (2021) interpret the negative relation between client trading activity and spreads as reflecting the strength of the trading relationship between dealer and client. However, trading relationships are relatively unimportant in the highly-liquid EUR-USD spot market, where all major dealing institutions maintain EUR inventory in nostro accounts at Eurozone banks, than in bond markets, where dealers often reach out to clients manage inventory before a trade with another client (Feldhütter and Poulsen, 2018). We therefore interpret the negative relation in terms of Bernhardt et al.'s (2005) theory of activity discounts.

[11] The FX core market was originally a pure interbank market using telephones. Reuters and EBS introduced interdealer matching engines around 1993. Around 2000 the interdealer platforms began to permit a few highly sophisticated clients to trade via prime brokerage contracts. New order-driven platforms for all-to-all client trading emerge frequently, though many do not survive.

From these data we select all spot trades and then exclude the following: the bank's internal trades, a handful of D2C trades on order-driven platforms or for which our bank is a price taker, and trades with other top-50 *Euromoney* dealing banks. The final sample includes 257,421 D2C trades worth €126.0 billion.

2.3.1 Clients

The bank has 2,288 active clients of seven types: hedge funds, client banks, brokers, real-money funds, private clients, MNCs, and SMEs. Among the hedge funds, roughly half engage in HFT. Most client banks are small and/or regional banks, many from countries with limited financial development; a few are sophisticated niche banks that sometimes trade engage in HFT. Most institutions in the bank's "broker" category are retail forex platforms or they act as agents for other institutions; a few are commodity trading advisors or large MNCs that trade actively. Real-money funds are low-leverage asset managers such as mutual funds, pension funds, insurance firms, or endowments. We include government entities in this category because their total trading is too limited for reliable statistical inference. Private clients manage the wealth of individuals or families. For the bank an MNC is defined as a non-financial firm with specialized treasury unit or at least €100 million in annual sales; an SME is defined as a non-financial firm with annual sales below €100 million and no treasury unit.

Table 2.1 provides descriptive statistics by client type. For a given descriptor we calculate each client's average and then take the mean across clients. The sample average trade size, €1.6 million, ranges widely: it is below €0.4 million for private clients and SMEs and above €4 million for hedge funds. The average client trades about 30 times per month, but this is skewed upwards by the brokers, whose monthly trades reach 470. Private clients, real-money funds, and non-financial clients typically trade about once per month.

The mean client-average effective spread is 32.9 pips or 25.1 basis points, with half-spreads measured as the distance from the D2C price to the contemporaneous Reuters mid-quote. To comparing execution costs across markets we need the mean spread across all trades, which is far lower at 1.2 pip or 0.95 bp. This is a bit below the mean effective spreads on S&P 500 stocks, 3.2 basis points (Hagströmer, 2021), and well below mean spreads in other OTC markets including FX derivatives, 16.4 bps (Hau et al., 2021), and US corporate bonds, 24.0 bps (Goldstein and Hotchkiss, 2020, round-trip trades that are exclusively D2C).

Variation in EUR-USD spreads across FX clients substantially exceeds available measures of variation across assets. The coefficient of variation for client-average half-spreads in spot EUR-USD, 1.3, is roughly twice the coefficient of variation for S&P 500 stocks, 0.7 (Hagströmer, 2021).

Average spreads in bps range across client types by almost two orders of magnitude, from 0.8 bps for hedge funds to 58.8 bps for private clients. The range in spreads across US corporate bonds with varied maturity (short, medium, long) and credit rating (AAA to C), is normally about one order of magnitude, from 11.3 bps for the safest short-maturity bonds to 105.3 bps for C-rated long-maturity bonds (Feldhütter and Poulson, 2018).

Spreads also vary widely across clients in other OTC markets. Mean client spreads in FX derivatives range from 2.5 pips to 30.1 pips for clients at the 25th and 75th percentiles, respectively (Hau et al., 2021). The wide variation of OTC bond spreads across trade sizes has been interpreted to represent differences across clients, as discussed in Section 2.2. Green et al. (2007) find that mean effective spreads for U.S. muni bonds range from 16 bps for large trades to 2300 bps for small trades. Goldstein et al. (2007) find that mean round-trip spreads per $100 face value of US corporate bonds range from $0.52 for large trades to $2.37 for small trades.

Our analysis focuses on the dealer choice variable, the markup, which ranges from 0.4 pip for hedge funds to 38.4 pips for private clients. The vast majority of markups are now set by an algorithm designed and parameterized by a dedicated team of financial engineers or "e-traders" with advice from the FX sales team. The remaining 1% of markups are set by an interbank dealer.

2.3.2 Trading Platforms

Clients of our bank request quotes on over 20 platforms that we partition into four groups: direct trades, single-bank platforms, multibank request-for-quote platforms, and automated programming interfaces. Table 2.1 presents mean platform shares, calculated as the share of a client's trades on a given platform type.

Direct trade (DT). A client requests quotes directly from a human dealer using telephone, email, txt, or another communication channel familiar from daily life. Quotes for about one third of direct trades are generated algorithmically; quotes for the remainder are selected by an interbank dealer. Direct trades have relatively high order-handling costs because they require the expensive time of a human dealer, unlike the other trading platforms. Private clients and SMEs rely on direct trading almost exclusively; MNCs also rely heavily on direct trading. Other clients rely little on direct trading.

Single-bank platform (SBP). A desktop user interface through which a client interacts electronically with a specific dealing bank. Over a simple SBP a client can instruct the bank to trade at a daily fix. On a sophisticated SBP a client can click-and-deal on live-streaming quotes. The sample includes 75,329 SBP trades. SBPs are in most frequent use among client banks (55% of trades, on average), real-money funds (30%), and MNCs (22%).

Multibank request-for-quote platform (MRFQ). An electronic platform over which clients request quotes from multiple dealing banks simultaneously.[12] Installing and maintaining MRFQ platforms requires non-trivial technical sophistication and these platforms are not available to institutions with just one or two FX dealing relationships. The sample includes 69,551 MRFQ trades. MRFQs are used most frequently by hedge funds (50% of trades), brokers (56%), and real-money funds (52%).

Application Programming Interface (API). Functionality embedded in some SBPs that enables clients to execute algorithmic and HFT trades. Substantial technical sophistication is required because clients must establish and maintain API connections and the trading algorithms. The sample includes 108,045 API trades. API connections are used most frequently by hedge funds (25% of trades on average) and retail FX brokers (23%) and rarely by other clients.

2.4 OTC Markups: Estimating Models

This section presents three nested markup decomposition models. The first two assume perfect competition; the third assumes that OTC dealers have market power.

2.4.1 Model 1: Perfect Competition and Dealer Costs

Our first model assumes, consistent with standard models of order-driven markets, that spreads are determined by order handling, inventory management, and adverse-selection costs (e.g., Glosten and Harris, 1988; Huang and Stoll, 1997).

Order-handling costs include dealer time, physical space, information resources, and trading technology. These fixed costs are highest when the trade involves a human interbank dealer or salesperson. We model the order-handling costs for trade t as follows:

$$OrderHandling_t = \varphi_0 + \varphi_1 Sz_t + DT_t\,(\varphi_2 + \varphi_3 Sz_t)\,. \qquad (2.1)$$

Sz_t is (log) trade size in EUR; DT_t is an indicator for direct trades. We expect φ_0, $\varphi_2 > 0$ to capture fixed costs. We expect φ_1, $\varphi_3 < 0$ because a given fixed cost can be covered by a smaller markup on a larger trade (Edwards et al., 2007).

Inventory management costs theoretically include both carry and risk, but costs of carry are generally negligible in FX because dealers infrequently hold positions overnight (Bjønnes and Rime, 2005). Inventory risk should rise with Sz_t and $Volatility_t$ (Stoll, 1978), market illiquidity (Bessembinder, 1994), and the time between trades (Tkatch and Kandel, 2008). We model the inventory risk of trade t as follows:

$$Inventory_t = \lambda_1 Sz_t + \lambda_2 Volatility_t + \lambda_3 RSprd_t + Time_t \lambda_4'. \qquad (2.2)$$

[12] The platform group also includes a few trades over voice brokers.

$Volatility_t$ is measured the high-low range of the previous trading hour, with mean and standard deviation 31.7 pips and 17.5 pips, respectively. The high-low range is chosen because it is always apparent to human dealers, but most results are robust to using realized volatility. *RSprd*, the core spread from Reuters in pips, captures illiquidity. Reuters' minimum spread was one pip during our sample period. *$Time_t$* is a vector of three time-of-day indicators to capture waiting time (Tkatch and Kandel, 2008). This is shortest during European hours (8-17 GMT), when EUR-USD trading is most active; longest during the market's brief overnight period (22-23 GMT); and intermediate during the New York afternoon (17-22 GMT). Asian hours are excluded. Standard theories predict $\lambda_1, \lambda_2, \lambda_3 > 0$, $\lambda_{4,ON} \geq \lambda_{4,NYC} \geq \lambda_{4,Europe} < 0$, $\lambda_{4,ON} > 0$. Liu, Makarov, and Schoar (2023).

In theory, inventory risk could also be influenced by client credit risk. However, FX dealing rooms do not manage this risk via bid-ask spreads, according to dealers and other market professionals. They rely instead on credit limits, prime brokerage contracts, or margins if a client's credit quality is difficult to ascertain.

Adverse Selection. OTC dealers, like all market makers, face adverse-selection risk when trading with informed clients. Because those dealers know their potential D2C counterparty when quoting, their response to information will be conditioned on the client's own tendency to be informed (Ramadorai, 2008). Dealers assess this tendency during long-term dealing relationships.

Following Anand and Subrahmanyam (2008) we measure information, *$Info_c$*, as the client's average signed one-minute post-trade return, with returns calculated using Reuters mid-quotes. (Our main are sustained when *$Info_c$* is measured using thirty-minute post-trade returns.) The mean of *$Info_c$* is statistically zero for all client types except hedge funds (Table 2.1). A hedge fund with *$Info_c$* one standard deviation above the hedge-fund mean, trading at the hedge-fund average frequency, would earn an extra 5.8% annually. We model adverse-selection costs as a linear function of client information:

$$Information_{c(t)} = Info_{c(t)}\pi \ , \ \pi > 0 \ . \qquad (2.3)$$

Model 1, in sum. Equations (2.1) through (2.3) imply the following model of OTC markups under perfect competition:

$$Markup = \varphi_0 + (\varphi_1 + \lambda_1)Sz + (\varphi_2 + \varphi_3 Sz)DT + \lambda_2 Volatility + \lambda_3 RSprd + Time\lambda_4' + Info\pi \ . \qquad (2.4)$$

2.4.2 Model 2: Perfect Competition, Costs and Future Profits

OTC dealers know their potential counterparty when quoting so they can rationally adjust markups in light of the potential profits from attracting a client's future business. The literature identifies two potentially profitable characteristics.

Client Trading Activity. Rational OTC dealers will quote narrower spreads to attract the future business of relatively active clients (Bernhardt et al., 2005). We assume a linear relation between this "activity discount" and client c's (log) daily average trading in EUR, ln($TrdVol_c$):

$$ActivityDiscount_{c(t)} = \beta\, ln(TrdVol_{c(t)}),\ \beta < 0\ . \qquad (2.5)$$

Client information. OTC dealers may find it profitable to narrow the spread to attract the future business of some informed clients, a strategy known as "information chasing" or "strategic dealing." Learning whether an informed client is buying or selling can help the dealer make profitable parallel trades in the core market (Naik et al., 1999) or wisely adjust their quotes to others (Pintér et al., 2022). If the anticipated profits from trading with an informed client match the anticipated adverse selection costs, or if a client is simply uninformed, a dealer can rationally adopt a third strategic option: make no information-motivated adjustment to spreads.

Model 2 allows dealers to vary their strategic response to information across client type:

$$Information_{c(t)} = Info_c \pi_{j(c(t))}\ ,\quad j = \{HF, ClBk, Broker, RM, Priv, MNC, SME\}. \qquad (2.6)$$

Information chasing implies $\pi_{j(c(t))}$<0; protecting the bank from adverse-selection risk implies $\pi_{j(cc(t))}$>0; the absence of any information adjustment implies $\pi_{j(c(t))}$=0.

Model 2, in sum. Model 2 comprises Model 1 (Equation (2.4)) plus Equations (2.5) and (2.6):

$$\begin{aligned} Markup = {} & \varphi_0+(\varphi_1+\lambda_1)Sz+(\varphi_2+\varphi_3 Sz)DT+\lambda_2 Volatility+\lambda_3 RSprd+Time\lambda_4' \\ & +\beta\, Ln(TrdVol) + \textstyle\sum_j InfoCT'_j\pi_j, \quad j=\{HF, ClBk, Broker, RMF, Priv, MNC, SME\}. \end{aligned} \qquad (2.7)$$

CT_j is an indicator variable for clients of type j.

2.4.3 Model 3: Market Power

Model 3 assumes that dealers have market power. As shown by Duffie et al. (2005), equilibrium OTC client spreads should be inversely related to a client's negotiating leverage or "outside options," which they parameterize as search costs. Green et al. (2007, p. 298) interpret negotiating leverage in terms of "sophistication" and highlight the importance of familiarity with the market: "we would expect more sophisticated customers to come to the market with a better understanding of the dealers' costs and resale opportunities and to therefore demand prices leading to lower dealer's profits." Sophistication in liquid OTC markets also involves the technical competence required to exploit the range of available trading technologies. Hendershott et al. (2020) and Hau et al. (2021) highlight another influence on a client's negotiating leverage, the number of its dealing relationships, which we label "potential counterparties." Not that the assumption that dealers have market power has no implications for client rationality. The rest of this section clarifies how Model 3 captures these client characteristics and provides evidence that the investments required to achieve sophistication and to secure multiple FX dealing relationships may not be cost-effective for certain clients.

Sophistication

Following Liu et al. (2023) we hypothesize that a client's platform choices carry information about its sophistication. Heavy reliance on direct trading, for example, suggests low sophistication because the client could instead trade over SBPs, which are low-cost, easy-to-install, and user-friendly. Clients with three or more dealing relationships can also trade over low-cost, user-friendly MRFQ platforms. We proxy a client's sophistication in part with 3x1 vector of its platform shares for direct trades, SBP trades, and MRFQ trades (APIs are omitted).

Sophistication requires market-specific knowledge plus the time and technical competence to adopt new trading methods. Institutions will invest in sophistication to the extent that they face the incentive to minimize execution costs, incentives that are stronger at institutions that trade more frequently, at larger firms, and at financial firms. If platform shares are influenced by sophistication they will also be strongly related to these incentives to invest in sophistication.

Trade frequency. Figure 2.1A shows that, for institutions that trade more than once per month, the share of direct trading is inversely related to trade frequency, consistent with the hypothesis that the inventive to invest in sophistication drives sophistication which, in turn, drives platform choice. Figures 2.1B and 2.1C provide further evidence for this hypothesis: Figure 2.1B shows that reliance on SBPs is positively related to trade frequency for financial clients and SMEs; Figure 2.1C shows that trade frequency is positively related to MRFQ reliance for MNCs.

For clients that trade just once per month – a group that includes 84% of SMEs, 97% of private clients, and 75% of MNCs – the absence of strong relations between trade frequency and platform choice is unsurprising. At a client trading once per month (week), the employee who trades will spend at least 99.5% (98.8%) of their time on other responsibilities such as executing payments, managing bank accounts, forecasting cash flows, and negotiating loan terms. Given that imbalance, those other responsibilities will determine compensation, not execution quality.

Client size. Small institutions may not reward execution quality because they are unable to monitor execution costs. SMEs are defined as having no treasury function and thus necessarily lack the expertise for transaction cost analysis (TCA). Nor can most SMEs justify outsourcing TCA to a specialist consulting firm: at the average SME half-spread, the execution cost of the mean SME trade (€350K) is just €1,138. This cost incurred monthly would amount to just 0.03% of revenue for an SME with €50 million in sales. Private wealth managers are similarly unlikely to have access to the expertise required for TCA. Consistent with these observations about the importance of size, 95.2% of SMEs (100.0% of private clients) rely on direct trading but 61.9% of MNCs (19.4% of other financial clients) do so. To test the statistical significance of these differences we note that clients make platform choices independently of other clients, so the number of clients choosing a given platform has a binomial distribution. We test the null that MNCs trade directly with probability 95.2%, like SMEs, and the null that SMEs trade directly with probability 61.9%, like MNCs. Both nulls are rejected at the 0.001% level. Two tests of the hypothesis that the probability of direct trading is the same for private clients and other financial clients are likewise rejected at the 0.001% level.

Financial vs. non-financial clients. Financial clients are generally more likely to reward execution quality than non-financial clients. At a hedge fund execution quality is implicitly rewarded through the dependence of bonuses on the firm's "2-and-20" compensation structure, and it may be explicitly rewarded as well. Many financial firms hire professional traders, who bring market knowledge to the institution and share it with others. Non-financial firms, including MNCs, generally do not reward execution quality. This difference suggests that financial firms will rely less on direct trading than non-financial firms, and the data support this inference. The share of clients that rely on direct trading is 25.5% for financial clients and 90.4% for non-financial clients. Symmetrically, 47.7% (46.2%) of financial clients rely on SBPs (MRFQs) but just 5.0% (5.2%) of non-financial clients. These differences all prove significant at the 0.001%.

Platform shares help capture a further influence of sophistication: the client's ability to recognize and avoid a classic dealer strategy for widening the markup. If the client normally buys (sells), the dealer quotes a conventional spread while shifting both quotes upward (downward) relative to the prevailing market price. This raises half-spread for a trade in the likely trade direction but can go unnoticed by unsophisticated clients because the spread is unremarkable. Model 3 captures this by including $Skew_c \equiv ShrDT_c BuySell_c$. $ShrDT_c$, the client's share of direct trading, serves as an inverse measure of sophistication. A client's tendency to trade in one direction is captured with $BuySell_c \equiv |\#Buys_c - \#Sells_c| / (\#Buys_c + \#Sales_c) \in [0,1]$.

We assume that $Soph_c$ and $Skew_c$ both influence markups linearly:

$$Sophistication_{c(t)} = Soph_{c(t)}\, \delta_2' + \theta\, Skew_{c(t)}\,. \qquad (2.8)$$

We expect $\delta_{2\text{-}Direct} > 0$, $\delta_{2\text{-}Direct} > \delta_{2\text{-}SBP}$, $\delta_{2\text{-}MRFQ}$, and $\theta > 0$.

Potential Counterparties

Clients with more potential counterparties have lower search costs and higher price elasticity for liquidity – that is, they are more likely to turn down unattractive quotes – so dealers will quote them narrower bid-ask spreads in equilibrium (Ho and Stoll, 1981; Duffie et al., 2005). FX clients with prime brokerage contracts have the most potential counterparties because they can trade with the entire professional trading community. At the other extreme, many clients cannot turn down a dealer's quotes because they have just one dealing relationship, or perhaps two. For small clients this constraint can reflect rational client decisions; for client banks the constraint is often imposed by the major dealing banks.

Clients that trade infrequently can rationally choose to have few dealing relationships because each one is costly to establish. Each client of a dealing bank must have a line of credit from that bank, and to quality the client must produce detailed financial accounts that demonstrate creditworthiness. Furthermore, they must do so not once but repeatedly over the years. Market contacts report that most private clients, most SMEs, and many small MNCs choose to have one or at most two FX dealing relationships for this reason. Evidence consistent with those reports comes from MRFQ platform shares. MRFQs typically require clients to have three or more dealing relationships, so these platform shares represent a lower bound on the share of unconstrained clients. Consistent with reports, just 5.3% of non-financial clients rely on MRFQs, far lower than the 50.4% share for financial clients; binomial tests confirm that the difference is highly significant. Further, SMEs and private clients – the smallest non-financial and financial clients – have MRFQ shares of 1.4% and 0.0%, respectively, well below the shares for MNCs and the remaining financial clients, 27.4% and 49.9%, respectively. Finally, Figures 2.1B and 2.1C support the hypothesis that these differences reflect a paucity of dealing relationships at smaller institutions: while MNCs shift strongly towards MRFQs as trade frequency increases, SMEs shift instead towards SBPs. (PCs trade at most once per week.)

The literature highlights two additional reasons why OTC clients might choose to have few dealing relationships, but these do not seem relevant to EUR-USD. First, clients can extract greater activity discounts by concentrating their trades with fewer dealers (O'Hara and Zhou, 2021). However, EUR-USD activity discounts are small according to Section 2.5 coefficient estimates. Second, clients with fewer dealing relationships may receive better service (Yueshen and Zou, 2023). However, the services in question, such as greater speed if the dealer has the asset in inventory, are relevant for bonds but not EUR-USD. The major dealers almost invariably hold EUR and USD and, if not, those currencies are always accessible in the core market.

Some banks from countries with low financial development ("LFD") are constrained to have just one FX dealing relationship. Major FX dealers are reluctant to trade with such banks due to the LFD countries' limited banking regulation, limited rule of law, and/or limited financial stability. These disadvantages are mitigated, however, when the LFD bank holds deposits at the major dealing bank as part of a correspondent banking relationship. According to a Treasury expert from our bank, these deposits can be used to settle FX trades, eliminating settlement risk. Further, they are typically stable and thus profitable for the dealing institution.

Figure 2.1D provides evidence that LFD banks are constrained from having many FX dealing relationships. We identify such banks as those with home country in the lowest 33% of 2012 IMF Financial Development Index: countries just below the 33.3% cutoff include Ecuador, Ghana, and Côte d'Ivoire. MRFQ shares are 8.6% for LFD banks and 47.7% for other banks, a difference that is significant at 0.001% and holds qualitatively for all trade frequencies.

Model 3 estimates the influence of a client's potential counterparties (*CP*) as follows:

$$CP_{c(t)} = \psi_{PB}PB_{c(t)} + \psi_{ClBk}ClBk_{LFD\text{-}NMc(t)} + \psi_{SME}SME_{NMc(t)} + \psi_{Priv}Priv_{c(t)}. \qquad (2.9)$$

PB is an indicator for clients that trade at least once with our bank over a prime brokerage contract. $ClBk_{LFD\text{-}NM}$ indicates LFD banks that do not trade over MRFQ platforms; SME_{NM} is defined similarly; *Priv* indicates private clients, none of which rely on MRFQ platforms. We expect $\psi_{PB} < 0$ and ψ_{ClBk}, ψ_{SME}, $\psi_{Priv} > 0$.

Intertemporal Smoothing

Dealers with market power may rationally smooth the bid-ask spread over time within client relationships. This should enhance client loyalty because customers in any setting tend to find frequent price changes "confusing, frustrating, and annoying" (Dholakia, 2016, p. 1). Smoothing can also clarify that a dealer's liquidity prices are competitive by eliminating noise. Finally, smoothing can enhance a dealer's reputation for reliability. In the survey of Cheung and Chinn (2001, p. 444), two-thirds of responding FX dealers report that the dominant influence on their quoted interdealer spreads is the spread's "conventional level"; trade profitability is secondary. Dealers explain that "the ability to consistently offer quotes with … conventional spreads in a hectic market is regarded as an essential characteristic of a market leader."

We model smoothing as follows:

$$Smooth_t = Time_t\kappa_1' + \kappa_2 Volatility_t + \kappa_3 RSprd_t. \qquad (2.10)$$

Coefficients on $Time_t$ should capture dealer efforts to offset intraday seasonals in core spreads, which are largest overnight and smallest during European hours (Ito and Hashimoto, 2006). Under smoothing $\kappa_{1Europe}>0$, $\kappa_{1ON}<0$, and κ_{1NYC} has ambiguous sign. FX core spreads respond positively to volatility (Bessembinder, 1994), so smoothing implies $\kappa_2<0$. The Reuters spread captures dealer responses to high-frequency changes in core spreads; smoothing implies $\kappa_3<0$.

Model 3, in sum. Model 3 combines Model 2 (Equation (2.7)) with Equations (2.8), (2.9), and (2.10):

$$\begin{aligned} Markup = {} & \varphi_0+(\varphi_1+\lambda_1)Sz + DT(\varphi_2+\varphi_3 Sz_t)+(\lambda_2+\kappa_2)Volatility+(\lambda_3+\kappa_3)RSprd+Time(\lambda_4'+\kappa_1') \\ & +\beta\, CTV + \textstyle\sum_j InfoCT'_j\pi_j \qquad (2.11) \\ & +Soph\delta_2' + \theta\, Skew + \psi_{PB}PB + \psi_{ClBk}\, ClBk_{LFD\text{-}NM} + \psi_{SME}SME_{NM} + \psi_{Priv}Priv.^{13} \end{aligned}$$

[13] Asymmetric information could also influence the trade-size coefficient if informed trades tend to be large (Easley and O'Hara, 1987). However, order-splitting is standard practice in FX.

High correlations among client characteristics could conceivably generate multicollinearity in this model: informed clients have strong incentives to invest in sophistication and to trade frequently. There are many uninformed clients in FX, however, which could moderate the associated correlations. We assess this possibility using bilateral correlations, with the client's share of direct trading, $ShrDT_c$, serving as a scalar measure of sophistication. The correlations are small and vary in sign: $\rho(Info_c, ShrDT_c) = 0.10$; $\rho(Info_c, TrdVol_c) = -0.00$; $\rho(ShrDT_c, TrdVol_c) = 0.27$.

2.5 OTC Markups: Model Estimates

This section presents the results of estimating markup decomposition Models 1, 2, and 3. Section 2.6, which follows, uses Model 3 to identify the dominant contributors to the variation of spreads across clients.

The models are estimated using OLS with Newey-West standard errors clustered by date. As shown in Table 2.2, all models have adjusted R^2 of 0.50 or higher, which is encouraging; the highest explanatory power in related studies is 0.33.[14] Though differences in explanatory power can have many sources, but these could well reflect the advantages of our data insofar as they include precise markups and permit comprehensive coverage of key client characteristics.

2.5.1 Model 1: Costs under Perfect Competition

Model 1, which assumes that spreads reflect dealing costs under perfect competition, provides evidence that the high order-handling costs of direct trades are passed on as higher markups, as predicted. The positive and significant coefficient on *DT* implies that a human dealer's time increases the fixed cost by 67.6 pips (52 bps), though this figure falls sharply in Model 3. Consistent with the logic that per-unit order-handling costs decline with trade size, the coefficient on *DTSz* is significantly negative. A trade of average size, €1.6 million, pays an extra 11.1 pips if handled directly.

Model 1's coefficients do not support standard theories of inventory management and adverse selection. Contrary to inventory theory, the volatility coefficient is significantly negative (Stoll, 1978), rather than positive; the core-spread coefficient is insignificant rather than positive; and the coefficients on time-of-day indicators are positive for European hours, when wait times are shortest, and negative for overnight hours, when wait times are longest (Tkatch and Kandel, 2008). Contrary to adverse-selection theory (Glosten and Milgrom, 1985), the $Info_c$ coefficient is statistically insignificant. The economic effects implied by these coefficients are small as well as unpre-

[14] Hau et al. (2021); Bernhardt et al. (2005); Harris and Piwowar (2006); Goldstein et al. (2007); O'Hara et al. (2018).

dicted. A one-standard-deviation rise in volatility reduces the markup by 0.003 pip. Markups during European trading hours are higher than overnight markups by just 0.004 pip.

The negative sign of Model 1's estimated coefficient for *Sz*, $\widehat{\varphi}_1$, is consistent with earlier studies of quote-driven markets (e.g., Reiss and Werner, 1996; Harris and Piwowar, 2006; Bessembinder et al. 2006; Green et al., 2007; Edwards et al., 2007). In order-driven markets a larger trade walks up the book and necessarily pays a wider spread. Negative trade-size coefficients have therefore been interpreted as evidence that *Sz* captures the influence of client trading volume (e.g., Bernhardt et al., 2005), client information (Pintér et al., 2022), or client sophistication (e.g., Green et al., 2007). We evaluate those claims with Models 2 and 3.

2.5.2 Model 2: Costs and Benefits Under Perfect Competition

Model 2 continues to assume perfect competition but incorporates the complementary hypothesis that dealers narrow spreads to attract clients whose business can provide future benefits. The results support this additional hypothesis. Markups are inversely related to client trading activity, as predicted in Bernhardt et al. (2005), though the effect is not large. A rise in activity from the SME mean (once per month) to the hedge fund mean (1.7 trades per day) reduces the markup by 1.3 pip. For brokers, markups are also inversely related to private information, consistent with Naik et al. (1999) and Pintér et al. (2022). A one-standard-deviation increase in broker information reduces their markup by 3.2 pips. This would reduce execution costs by 1.1% annually for a broker with mean trade size and mean trade frequency.

Model 2 also provides evidence that OTC dealers exploit their strategic flexibility with respect to client information. Though the negative $Info_c$ coefficient for brokers points to information chasing, $Info_c$ coefficients for hedge funds and client banks are positive and significant, consistent with adverse selection (Glosten and Milgrom, 1985; Easley and O'Hara, 1987). A one-standard-deviation increase in $Info_c$ for a hedge fund (client bank) raises its markup by about half, or 0.25 pip (0.39 pip).

The $Info_c$ coefficients are insignificant for non-financial clients, real-money funds, and private clients. This could, in theory, reflect a balance between adverse-selection costs and dealing benefits. However, the descriptive statistics of Table 2.1 suggest that these clients are simply uninformed, a hypothesis that is supported in other studies (Menkhoff et al., 2016; Bjønnes et al., 2021; Ranaldo and Somogyi, 2021) and consistent with known trading incentives. At non-financial firms there is generally no expectation of returns from FX trading. They use a foreign currency primarily as a medium of exchange, rather than a store of value, and internal governance mandates treat it as a cost within the Finance function. Indeed, the bylaws of many non-financial firms discourage or prohibit speculative trading because it brings rogue-trader risk, which is costly to manage. Actively-managed real-money funds and private clients use currency as a store of value, which could motivate investments in information, but index funds have no such motivation. Industry experts report that most such funds treat FX trading as an unavoidable administrative cost (Taylor and Farstrup, 2006; Galanek, 2010).

Model 2's volatility coefficient and time-of-day coefficients continue to be inconsistent with inventory theory. The trade-size coefficient continues to be negative and its magnitude is essentially unchanged, which suggests that the negative coefficient in Model 1 was not due to the model's exclusion of client trading activity or information.

2.5.3 Model 3: Market Power

Model 3 provides strong evidence that OTC dealers have market power. Platform-share coefficients are highly significant and conform to predictions: positive for direct trading and smaller (indeed, negative) for SBPs and MRFQs. According to these coefficients the difference in sophistication between an SME and a hedge fund raises SME markups on all trades by 3.4 pips. Further evidence for dealer market power comes from *Skew*. A client at the mean for both directional predictability ($Buy/Sell_c$) and the share of direct trading ($ShrDT_c$) pays a markup that is 0.8 pip wider than a client with no directional predictability ($Skew_c$=$Buy/Sell_c$=0).

Model 3 also supports the hypothesis that a client's negotiating leverage is inversely related to the number of its potential counterparties. Clients with prime broker arrangements – and thus the most potential counterparties – pay markups that are lower by 0.4 pip, ceteris paribus; LFD client banks pay an extra 0.7 pip; SMEs without access to MRFQ platforms pay an extra 10.7 pips; and private clients pay an extra 23.6 pips.

Model 3's coefficients also provide evidence consistent with intertemporal smoothing, another manifestation of dealer market power, and can reconcile the model with standard theories of dealer costs. The negative and significant volatility coefficient supports the smoothing hypothesis and indicates that smoothing ($\kappa_2 < 0$) dominates inventory risk ($\lambda_2 > 0$) with respect to volatility. The continued insignificance of the core-market half-spread is consistent with the hypothesis that its impact through smoothing, $\kappa_3 < 0$, and its impact through market illiquidity, $\lambda_3 > 0$, have similar magnitudes. Coefficients on time-of-day indicators retain their signs and significance from Models 1 and 2 – positive for European hours and negative during overnight hours – and the associated difference in markups between European and overnight hours remains below 0.01 pip. This pattern supports the smoothing hypothesis and indicates further that smoothing dominates inventory risk (λ_4) with respect to intraday seasonals in FX spreads.

Model 3's estimate for the fixed cost of a direct trade, 30.7 pips, is just half the estimate from Model 1. An average-sized trade (€1.6 million) is now estimated to pay an extra 1.2 pips when handled directly, far below the 11.1-pip and 10.0-pip estimates from Models 1 and 2, respectively. This change reveals that direct trading influences markups in multiple ways, only one of which is appropriately captured by order-handling costs. Direct trading also signals a lack of sophistication to a dealer, who responds by quoting wider spreads and skewing the quotes for clients with predictable trade direction.

With the regression expanded to include sophistication and potential counterparties, the coefficient on *Sz* becomes statistically insignificant. This implies that its influence through inventory risk ($\lambda_1 > 0$) roughly matches its influence through order-handling costs ($\varphi_1 < 0$). It also suggests that Green et al. (2007) were correct to interpret negative trade-size coefficients for OTC spreads as reflecting the omitted variable of client negotiating leverage.

2.5.4 Alternative Hypotheses

We consider two alternative hypotheses regarding the high spreads paid by clients with low sophistication or few potential counterparties.

Alternative 1: *The market is fully competitive and efficient and the price of liquidity only reflects dealer costs. The clients who seem less sophisticated pay more for FX liquidity because they consume more of the dealing room's unpriced services, such as economic research, technical trading insights, or advice on hedging.* The evidence presented so far suggests, the opposite pattern. Clients with strong incentives to minimize execution costs – those that trade more frequently, are larger, and are in finance – would consume more of these unpriced services and would pay wider spreads, other things equal. To investigate, we interviewed the head of e-trading at our bank plus other FX-market professionals. They consistently report that SMEs and private clients are rarely interested in market color from the bank's sales team, and are similarly disinterested in the bank's economic research or technical trading insights. Hedge funds, by contrast, make a point of gathering perspectives from many banks.[15] In short, client consumption of unpriced dealer services cannot explain cross-client variation in OTC spreads.

Alternative 2: *The market is fully competitive and efficient and the price of liquidity is adjusted to support other client business within the bank.* Bonuses for dealers and salespersons are based on total dealing-room profits from trading a given currency, which are calculated daily and reviewed frequently with the head trader. This encourages dealers and salespersons to support each other but provides no incentive for them to support other bank functions. According to our interloc-

[15] After our sample period European authorities prohibited banks from bundling research and trading services.

utors from the market, FX dealing operations make little effort to support other client businesses with their bank.

To market participants, the influence of dealer market power in FX is unremarkable and has no moral valence. In their world it is common knowledge that dealers are ultimately responsible to shareholders, that FX dealing banks always trade as principals, that a dealer's job is to maximize trading income, and that FX dealers therefore maximize client spreads. FX dealers are astute: they appreciate the importance of relationships and will ingratiate themselves with clients by engaging in small talk, travelling to far-away HQs to assess client needs, and hosting gala events. But these astute dealers also run experiments – and willingly lose a few trades – to identify precisely how much those now-friendly clients will pay for liquidity. The extent to which dealers widen spreads when circumstances permit is highlighted in Osler and Savaser's (2022) study of FX dealing at global custody banks, where opacity is extreme. Those dealers often set client purchase (sale) prices just inside each day's high (low), achieving effective spreads nearly as wide as the day's trading range.

2.5.5 Robustness Tests

Model 3's qualitative conclusions about OTC spreads are robust, as shown in Table 2.3. Robustness test 1 replaces the original $Info_c$, average one-minute post-trade returns, with average 30-minute post-trade returns. The model's explanatory power and most coefficients are little changed, so the model's qualitative conclusions are sustained. The coefficients on $Info_c$ decline in absolute magnitude for informed clients and the coefficient for client banks becomes insignificant. On balance, however, 30-minute $Info_c$ has a more substantial effect on markups than 1-minute $Info_c$ because 30-minute post-trade returns tend to be larger. A one-standard-deviation rise in a hedge fund's 30-min $Info_c$ now raises the markup by **2.4** pips, far more than the 0.5-pip rise for 1-minute $Info_c$.

Robustness test 2 replaces the measure of volatility based on high-low ranges with the realized volatility of one-minute mid-quote Reuters returns over the same 60-minute pre-trade interval. Volatility itself becomes insignificant, consistent with the hyothesis that perceived volatility is more strongly influenced by the visible trading range than a constructed measure. The remaining coefficients are qualitatively unchanged.

Robustness 3 captures client sophistication with the client's direct-trade platform share, $ShrDT_c$, plus a single indicator variable for two groups of FX experts: hedge funds and client banks. The coefficient on $ShrDT_c$ is essentially unchanged and the new indicator's coefficient is significantly negative, as expected. Otherwise the model's explanatory power is stable and the remaining coefficients are little changed.

Robustness 4 limits LFD countries to those with Financial Development Index in the lowest 20%. This brings a significant increase in the LFD-bank coefficient, from 0.77 to 0.97, as one would expect, but leaves the other results qualitatively unchanged.

Robustness test 5 allows the markup to differ according to client trade direction by introducing a dummy for EUR sales. This variable's coefficient is significantly negative but has tiny economic magnitude: margins on an EUR sale are smaller by 0.08 pip. The model's explanatory power is stable and other coefficients remain little changed.

Cross-Client Variation in OTC Liquidity Prices

We finish our analysis of OTC client liquidity prices by identifying the dominant sources of cross-client variation. Why do some OTC clients pay spreads of 75 pips while others pay less than one pip? We focus this analysis on markups, rather than spreads, because at a given moment every client spread embeds the same core spread. In consequence, the cross-client variance of markups accounts for over 99% of the cross-client variance of spreads.

Robustness test 6 considers the 3,066 trades priced by interbank dealers rather than the bank's e-trading algorithm. This can occur because a special client reaches out directly to the interbank dealers, because a trade is worth €25 million or more, or because the trade otherwise appears toxic to the bank. These trades tend to be placed by hedge funds, client banks, and brokers (76%); 43% are initiated via direct communication with dealers, 22% are initiated via SBPs, and 45% are initiated via MRFQs. They tend to be larger than average, accounting for 15% of total traded value, though many are small. With this subsample the model's explanatory power falls to 0.38, though that remains strong relative to existing analyses of OTC spreads (Bernhardt et al., 2005; Harris and Piwowar, 2006; Goldstein et al., 2007; O'Hara et al., 2018; Hau et al., 2021). The qualitative conclusions from the full-sample regression are generally sustained but the toxic nature of these trades brings two noteworthy changes. First, the relation between trade size and markup becomes positive rather than insignificant, implying that the influence of inventory risk, $\lambda_1 > 0$, dominates the influence of order-handling costs, $\varphi_1 < 0$. Second, dealer responses to private information reveal much greater concern for adverse selection. The response to hedge-fund information rises by a multiple of six; the response to client-bank information increases by an order of magnitude to the roughly same level as the new hedge-fund response; the response to broker private information goes from moderately negative to strongly positive. The annualized cost of a one-standard-deviation rise in client information, with group-average trading frequency, becomes 18.1%, 18.0%, and 18.3% for hedge funds, client banks, and brokers, respectively.

2.6 Cross-client Variation in OTC Liquidity Prices

We finish our analysis of OTC client liquidity prices by identifying the dominant sources of cross-client variation. Why do some OTC clients pay spreads of 75 pips while others pay less than one pip? We focus this analysis on markups, rather than spreads, because at a given moment every client spread embeds the same core spread. In consequence, the cross-client variance of markups accounts for over 99% of the cross-client variance of spreads.

2.6.1 Cross-client Variation in Liquidity Prices

Our variance decomposition analysis begins with the fitted markup from Model 3, denoted $\widetilde{Markup}_t$, for each trade *t*. From these we calculate client-specific averages, $\widetilde{Markup}_c$, and the sample-wide cross-client variance, $\sigma^2_{\widetilde{Markup}_c}$. We use the same estimated coefficients to calculate fitted values of key markup components in versions of Equations (2.1), (2.2), (2.5), (2.6), (2.8), (2.9), and (2.10); averages for individual clients; and variances across clients. Variables for which theory predicts conflicting effects are included according to the sign of their estimated coefficients.[16]

Order-handling costs (*OH*): $\widetilde{OH}_t = \hat{\varphi}_0 + (\hat{\varphi}_1 + \hat{\lambda}_1) Sz_t + DT_t(\hat{\varphi}_2 + \hat{\varphi}_3 Sz_t)$, (2.1′)

Inventory risk (*INV*): $\widetilde{INV}_t = (\hat{\lambda}_3 + \hat{\kappa}_3) RSprd_t$, (2.2′)

Activity discount (*AD*): $\widetilde{AD}_{c(t)} = \hat{\beta} Ln(TrdVol_{c(t)})$, (2.5′)

Information (*INFM*): $\widetilde{INFM}_{c(t)} = \hat{\pi}_{j(c)} Info_{c(t)}$, (2.6′)

Sophistication (*SPH*): $\widetilde{SPH}_{c(t)} = Soph_{c(t)} \hat{\delta}_2' + \hat{\theta} Skew_{c(t)}$, (2.8′)

Counterparties (*CP*): $\widetilde{CP}_{c(t)} = \hat{\psi}_{PB} PB_{c(t)} + \hat{\psi}_{ClBk}\, ClBk_{LFD\text{-}NM\ c(t)} + \hat{\psi}_{SME} SME_{NM\ c(t)} + \hat{\psi}_{Priv} Priv_{c(t)}$, (2.9′)

Smoothing (*SM*): $\widetilde{SM}_t = Time_t \hat{\kappa}_1' + \hat{\kappa}_2' Volatility_t$. (2.10′)

Each component's proportionate contributions to cross-client markup variance is identified using the definition of variance, as shown below. Results are presented in Table 2.4.

$$1 = \sum_i \frac{\sigma_i^2}{\sigma^2_{\widetilde{Mkp}}} + \sum_i \sum_{j \neq i} \frac{Cov(\hat{\imath}, \hat{\jmath})}{\sigma^2_{\widetilde{Mkp}}}, \qquad i = \{OH, INV, AD, INFM, SPH, CP, SM\}. \qquad (2.12)$$

The contribution of a component's own variance, $\sigma_i^2 / \sigma^2_{\widetilde{Mkp}}$, is termed its "outright" contribution.

Cross-client markup variance is dominated by cross-client differences in dealer market power and order-handling costs. The outright contributions of client characteristics that cede market power to dealers – a lack of sophistication, predictable trade direction, and a dearth of potential

[16] For example, *Sz* is included in order-handling costs because its coefficient has negative sign; if the sign were positive it would be included in inventory risk.

counterparties – sum to 24.7%. Sophistication and potential counterparties are strongly correlated ($r(Soph_c,CP_c)$=0.65), so their covariance contributes a further 12.2% of variance. The minimum contribution of market power characteristics is thus 36.9%.

The outright contribution of order-handling costs is also large, at 18.0%. Market power and order-handling costs make indirect contributions through their strong positive correlations: a client who tends to trade directly signals low sophistication to the dealer (correlation($\widetilde{OH}_c$,$\widetilde{SPH}_c$) =0.65) and tends to have few counterparties (correlation($\widetilde{OH}_c$,$\widetilde{CP}_c$)=0.74). Covariances between order-handling costs and market power variables jointly contribute 39.7% of variance. In total, dealer market power and order-handling costs 92.5% of cross-client variance in OTC markups.

None of the remaining determinants of dealing costs and benefits – inventory risk, smoothing, activity discounts, and client information – make outright contributions in excess of 0.3% of cross-client variation in markups, and their mutual covariances contribute only 2.1%. The small contribution of inventory risk to cross-client variance is unsurprising because it is largely determined by market-wide factors that affect all clients similarly. The unimportance of smoothing is unsurprising given the small economic effects of *Volatility* and *Time*.

This variance decomposition brings into focus the importance of client behavior for OTC liquidity prices. Variation of spreads across clients would shrink by over 90% if all clients invested in market knowledge and honed their negotiating skills, established many FX dealing relationships, and rewarded employees for minimizing execution costs so they avoided direct trading. The failure of clients to adopt these behaviors may appear inconsistent with rationality, and – as one would expect given the Dunning-Kruger effect (Kruger and Dunning, 1999) – every dealer has stories of clients who overestimate their trading acumen. Nonetheless, aggressive measures to minimize execution costs may simply not be cost-effective for the 97% of private clients and the 72% of non-financial clients that trade at most once per month. Together, these groups account for 44% of clients in our sample.

2.6.2 Profit-seeking vs. Utilitarian Clients

To further analyze the sources of cross-client spread variation we carry out the variance decomposition separately for two groups identified by Harris (2003): *profit-seeking* clients, a group that here comprises hedge funds, client banks, and brokers; and *utilitarian* clients, which here comprises real-money funds, private clients, and non-financial clients (Table 2.4, Columns 2 and 3). The absolute variance of markups is two orders of magnitude smaller for profit-seeking than utilitarian clients, at 4.8 pips^2 and 140.0 pips^2, respectively. It appears that most financial clients avoid direct trading, negotiate astutely, and have many FX dealing relationships. Nonetheless, the total contributions of order-handling costs and dealer market power are similar for the two groups: 89.2% for profit-seeking clients and 92.9% for utilitarian clients.

For utilitarian clients the outright contributions of order-handling costs and dealer market power are similar to those identified for the full sample. For profit-seeking clients, by contrast, the outright contributions of order-handling costs (26.2%) and sophistication (33.4%) are higher and the outright contribution of potential counterparties is lower (0.9%). The contribution of order-handling costs among these clients may reflect variation in trade toxicity, since toxic trades are more aggressively priced, rather than the factors highlighted for the full sample: preference for convenience, low negotiating skills, and a paucity of counterparties.

2.6.3 Robustness

We examine the robustness of these qualitative conclusions using four of the robustness tests reported in Section 2.5: robustness 1, which replaces one-minute with 30-minute post-trade returns in $Info_c$; robustness 3, which measures sophistication differently; robustness 4, which defines LFD banks more narrowly; and robustness 6, which constrains the sample to trades priced by a human interbank dealer. Results are provided in Table 2.5.

None of these alternatives suggest any changes to the qualitative conclusions from baseline estimates of Model 3. Client choices and constraints, as manifest in sophistication, counterparties, and order-handling costs, account for at least 83% of total cross-client variation. These shares are slightly lower for profit-seeking than utilitarian clients but never fall below 89%. Market power variables still account for lower shares of total variation for profit-seeking clients, but the differences are modest: market power accounts for 34% to 43% of total variance for profit-seeking clients and from 42% to 54% for utilitarian clients.

2.7 Concluding Remarks

This study provides a comprehensive analysis of variation in bid-ask spreads across OTC clients. We exploit the complete record of D2C trades by a major dealing bank in EUR-USD, a single contract for which daily volume exceeds daily volume in all US and European equities plus all US corporate and municipal bonds. These data have rare but important advantages for the study of OTC spreads: identifiers for individual clients and many client types, identifiers for over twenty trading platforms, and precise markups over the core-market (interbank) price.

EUR-USD spreads vary widely across clients. Spreads for the average SME or private wealth manager are an order of magnitude wider than spreads for the average hedge fund and the coefficient of variation for these OTC spreads is an order of magnitude larger than the coefficient of variation across S&P stocks (Hagströmer, 2021). Client spreads in FX forwards, which are likewise traded OTC, also vary widely (Hau et al., 2021).

The paper estimates three nested decomposition models tailored to liquid OTC markets. The model's dependent variable is the trade markup, which is superior to the effective spread for two reasons because it is the dealers' choice variable and because it is chosen after the dealer observes the other component of the client half-spread, the core-market half-spread. Model 1 assumes that spreads reflect dealer costs (Model 1) under perfect competition, where costs are higher for "direct" trades, which involve a human dealer, than for trades over electronic platforms. Model 2 assumes perfect competition but includes the possibility the dealers benefit from trading with active and informed clients (Bernhardt et al., 2005; Naik et al., 1999; Pinter et al., 2022). The third model assumes that dealers have market power relative to clients with limited negotiating skills or few potential counterparties (Duffie et al., 2005; Green et al., 2007).

The results indicate that dealer market power is the largest single contributor to cross-client variation in markups and thus spreads, accounting for at least 37%. The order-handling costs of direct trades are also important, contributing 16% outright. Market-power variables tend to be highest for clients that trade directly and their correlations jointly contribute a further 56%. These broad qualitative conclusions are sustained over numerous robustness tests and when clients are partitioned into profit-seeking and utilitarian clients, following Harris (2003).

Our analysis indicates that most cross-client variation would disappear if all clients fully exploited the available trading options, operated with maximum negotiating acumen, and maximized competition among dealers. Clients that pay high execution costs could be fully rational, nonetheless, achieving those standards is costly. It is costly to carry out the TCA required to reward employees for trading on low-cost platforms; it is costly to invest in trader negotiating skills; it is costly to establish and maintain new trading platforms; and it is costly to establish and maintain multiple dealing relationships. For the 40%-plus of FX clients in our sample that trade at most once per month, these costs could well exceed the associated savings.

Future research could productively examine the extent to which these findings apply in other OTC markets. Researchers could also fruitfully examine whether cross-client differences in OTC spreads have contributed to shifting financial market structures in recent decades.

Table 2.1:Descriptive Statistics

		All Clients	Hedge funds	Client Banks	LFD Banks	Brokers	Real-money	Private Clients	MNC	SME
A. Basic Descriptors										
Number clients		2,388	43	631	100	110	257	88	169	990
Number trades		257,241	6,624	62,239	11,491	170,668	1,185	113	1,261	3,660
Mean trades/client		108	154	99	115	1,552	5	1	7	4
B. Client-average Markups (pips)										
	Mean	16.34	0.41	0.62	4.14	1.31	3.75	38.36	9.68	32.38
	Max	239.29	8.78	60.00	239.29	46.00	66.41	102.00	120.00	104.00
	St. Dev	*21.18*	*1.69*	*3.93*	*23.92*	*6.54*	*10.25*	*17.83*	*17.33*	*18.58*
C. Client-average Trade size										
	Mean (€ mn)	1.57	4.72	2.63	0.54	1.11	3.70	0.17	2.43	0.35
	Max (€ mn)	170.00	22.16	170.00	8.25	24.00	100.13	2.43	38.15	25.00
	St. Dev. (€ mn)	*6.44*	*6.07*	*8.43*	*1.17*	*2.75*	*12.67*	*0.30*	*5.25*	*1.25*
	Mean (ln €)	12.16	14.37	13.26	11.93	12.45	12.68	11.19	12.73	11.21
	St. Dev. (ln €)	*2.13*	*1.44*	*1.75*	*1.48*	*1.54*	*2.47*	*1.42*	*2.56*	*1.85*
D. Client-average Daily Trading										
	Mean (€ mn)	0.79	2.73	1.44	0.47	6.56	0.14	0.00	0.28	0.01
	Median (€ mn)	0.01	0.40	0.25	0.07	0.95	0.01	0.00	0.02	0.00
	St. Dev. (€ mn)	*4.45*	*4.63*	*4.14*	*1.29*	*16.63*	*0.41*	*0.01*	*1.09*	*0.05*
	Mean (ln €)	9.63	13.28	12.29	10.90	13.25	9.17	7.14	9.48	7.60
	St. Dev. (ln €)	*3.18*	*2.04*	*2.27*	*2.34*	*2.88*	*2.76*	*1.46*	*3.05*	*2.06*
E. Client-average platform shares (%)										
Direct	Mean	52.76	16.01	12.75	5.91	5.01	17.39	100.00	53.05	94.80
	St. Dev.	49.05	34.97	30.87	*22.00*	*19.08*	*36.67*	*0.00*	*47.75*	*21.84*
SBP	Mean	24.03	9.20	48.33	94.09	16.00	30.37	0.00	22.27	3.77
	St. Dev.	*41.35*	*28.55*	*46.84*	*22.00*	*34.88*	*44.88*	*0.00*	*40.48*	*18.78*
MRFQ	Mean	21.22	49.55	37.42	0.00	55.80	52.22	0.00	24.68	1.22
	St. Dev.	*39.47*	*48.64*	*45.05*	*0.00*	*47.63*	*49.53*	*0.00*	*41.79*	*10.74*
API	Mean	2.00	25.24	1.49	0.00	23.19	0.02	0.00	0.00	0.20
	St. Dev.	*13.73*	*43.42*	*11.88*	*0.00*	*40.96*	*0.30*	*0.00*	*0.00*	*4.49*
F. Client Information (pips)										
	Mean	0.01	0.10	0.00	-0.16	0.07	-0.14	-0.31	0.03	0.10
	Median	0.00	0.16	0.01	-0.11	0.09	0.00	-0.14	0.00	0.04
	St. Dev.	*1.58*	*0.95*	*0.99*	*1.12*	*0.76*	1.67	*1.97*	*1.54*	*1.91*
G. Directional Predictability (Buy/Sell $\in[0,1]$**)**										
	Mean	0.32	0.33	0.32	0.40	0.17	0.24	0.09	0.31	0.36
	St. Dev.	*0.41*	*0.36*	*0.33*	*0.39*	*0.26*	*0.40*	*0.28*	*0.43*	*0.47*

Table shows summary statistics for markups measured in pips, where one pip is $0.0001/€. Data include all market-making EUR-USD trades by a top-10 EUR-USD dealing bank with non-dealing-bank clients during the 68 trading days from 2 January through 20 April, 2012. Direct: Client reaches out to an interbank dealer or salesperson by telephone, email, txt, etc. SBP: Single-bank platforms provide prices on the client's desk-top and connects directly to the dealing bank. API: Automated programming interface. MRFQ: Request for quote systems, one client requests near-simultaneous quotes from multiple dealing banks. Information: A client's average one-minute post-trade return in pips, where one pip is $0.0001/€; prices are core-market mid-quotes.

Table 2.1: Determinants of FX Client Markups

	Model 1: Perfect Competition, Costs Only	Model 2: Perfect Competition, Costs and Future Benefits	Model 3: Costs, Future Benefits, and Market Power
Inventory and Order-handling Costs			
Constant	0.321†	4.567†	2.176†
Sz (e-2)	-0.025†	-0.021†	-0.304
DT	67.628†	62.934†	30.706†
DT x Sz	-3.962†	-3.707†	-2.065†
Volatility (e-2)	-0.016†	-0.041	-0.074**
CoreSprd (e-2)	0.668	0.697	0.678
Time: Europe	0.169†	0.151†	0.125†
NYC	0.070†	0.103†	0.106†
Overnight	-0.248†	-0.134***	-0.043*
Information			
$Info_c$	0.013		
x Hedge fund		0.258†	0.540†
x Other Cl. Bks		0.392†	0.297***
x Broker		-0.873†	-0.384†
x Real-money		0.248	0.343*
x Priv. client		-2.480*	-1.194
x MNC		-0.448	-0.707
x SME		-0.405	-0.321
Activity Discount		-0.266†	-0.131†
Sophistication			
Shr Direct (e-2)			4.143†
Skew (e-2)			4.905†
Shr SBP (e-2)			-0.113†
Shr MRFQ (e-2)			-0.157†
Potential Counterparties			
x PB			-0.370†
x $ClBk_{LFD\text{-}NM}$			0.627†
x SME_{NM}			10.781†
x Priv			23.703†
Adj. R^2	0.499	0.513	0.580
N. Observations	257,241	257,241	257,241

Table reports coefficients from estimating Equation (2.11):

$$Markup = \varphi_0+(\varphi_1+\lambda_1)Sz + DT(\varphi_2+\varphi_3 Sz_t)+(\lambda_2+\kappa_2)Volatility+(\lambda_3+\kappa_3)RSprd+Time(\lambda_4'+\kappa_1')$$
$$+\beta\, CTV + \sum_j InfoCT'_j\pi_j$$
$$+Soph\delta_2'+\theta Skew+\psi_{PB}PB+\psi_{ClBk}\, ClBk_{LFD\text{-}NM} + \psi_{SME}SME_{NM}+\psi_{Priv}Priv.$$

Markup = D(pclient-pcore), Dt=1 (-1) for client purchases (sales). Sz: log trade size in EUR. DT: direct-trade indicator. Volatility: (log) Reuters high-low range over the previous hour. RSprd: core spread from Reuters. Time: indicators for trades during European hours, NYC hours, and overnight hours (Asian hours excluded). CTV: client's log trading volume. Info: client's average 1-minute post-trade return. Soph: client's shares of trades handled directly, over SBPs, or over MRFQs. Skew: client's direct-trading share times tendency to trade in one direction. PB: indicator for clients with prime brokerage contracts. ClBkLFD-NM: indicator for client banks from low-financial-development countries that do not trade over MRFQs. SMENM: indicator for SMEs that do not trade over MRFQs. Data: all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012. Robust standard errors clustered by date. Significance at 10%, 5%, 1%, and 0.1% levels indicated by *, **, ***, and †, respectively.

Table 2.3: Determinants of FX Client Markups: Robustness Tests

	Alt. *Info*	Alt. *Vola-tility*	Alt. *Soph-ist.*	Alt. LFD Client Bank	Client Sale Dummy	Priced by Interbank Dealer
	1	2	3	4	5	6
Costs						
Constant	2.036†	2.145†	2.629†	2.202†	2.209†	-5.131*
Sz (e-2)	0.103	-0.220	1.532	-0.418**	-0.251	0.541**
DT	32.315†	32.262†	32.008†	32.193†	32.220†	38.549†
DT x Sz	-2.085†	-2.081†	-2.069†	-2.078†	-2.079†	-2.708†
Volatility (e-2)	-0.085**	-0.039	-0.080**	-0.007**	-0.071**	-2.246*
CoreSprd (e-2)	0.674	0.699	0.661	0.733	0.666	7.389
Time: Europe	0.120†	0.112†	0.120†	0.123†	0.120†	2.340†
NYC	0.094†	-0.037	-0.099†	0.105†	0.105†	2.334
Overnight	-0.048*	-0.360†	-0.183†	-0.035	-0.0501**	-3.018†
Client sale					-0.078†	
Information						
x Hedge fund	0.285†	0.526†	0.775†	0.533†	0.537†	3.020†
x Client bk	0.003	0.303***	0.344***	0.300***	0.299***	2.878***
x Broker	-0.077**	-0.386†	-0.672†	-0.396†	-0.386†	2.583†
x Real-money	-0.049	0.356*	0.349*	0.357*	0.356*	0.249
x Priv. client	-0.096	-1.196	-1.196	-1.196	-1.195	-2.188
x MNC	-0.156	-0.751	-0.722	-0.740	-0.741	1.320
x SME	-0.046	0.320	0.345	0.320	0.321	-0.620
Activity Discount	-0.125†	-0.130†	-0.169†	-0.133†	-0.131†	-0.211†
Sophistication						
Shr Direct (e-2)	2.897†	2.830†	2.984†	2.836†	2.283†	3.524**
Skew (e-2)	4.807†	4.829†	4.799†	4.813†	4.840†	12.231
Shr SBP (e-2)	-0.106†	-0.112†		-0.117***	-0.116***	-2.102*
Shr MRFQ (e-2)	-0.165†	-0.155†		-0.156†	-0.155†	-0.561
HF + ClBks			-0.390†			
Constraints						
x PB	-0.391†	-0.360†	-0.183	-0.361†	-0.363†	-2.328***
x $ClBk_{LFD\text{-}NM}$	0.616†	0.628†	0.770†	0.967†	0.626*	29.113†
x SME_{NM}	10.790†	10.768†	10.438†	10.776†	10.798†	7.202**
x Priv	24 .051†	23.657†	23.254†	23.664†	23.675†	11.667**
Adj. R^2	0.580	0.580	0.581	0.581	0.580	0.378
N. Observations	257,241	257,212	257,241	257,241	257,241	3,066

Table reports coefficients from estimating Equation (2.11) using OLS with Newey-West …

$$Markup = \varphi_0+(\varphi_1+\lambda_1)Sz + DT(\varphi_3+\varphi_4 Sz_t)+(\lambda_2+\kappa_2)Volatility+(\lambda_3+\kappa_3)RSprd+Time(\lambda_4'+\kappa_1')$$
$$+\beta\, CTV + \sum_j InfoCT'_j \pi_j$$
$$+Soph\delta_2' + \theta Skew + \psi_{PB}PB + \psi_{ClBk}\, ClBk_{LFD\text{-}NM} + \psi_{SME}SME_{NM} + \psi_{Priv}Priv.$$

Regression variable are defined in the notes to Table 2. *Alt. Info*: $Info_c$ with 30-minute post-trade returns. *Alt. Volatility*: realized volatility. *Alt. Sophist*: SBP and MRFQ platform shares replaced by client-type indicators for hedge funds and client banks. *Alt. LFD Client Bank*: Low financial development defined as IMF Financial Development Index below the 20th percentile. *Client Sale Dummy:* Includes indicator for client sales. Data comprise all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012. Robust standard errors clustered by date. Significance at 10%, 5%, 1%, and 0.1% levels indicated by *, **, ***, and †, respectively.

Table 2.4. Sources of Cross-client Variation in OTC Effective Spreads

	All Clients	Profit-seeking	Utilitarian
	1	2	3
Cross-client variance, fitted markups ($pips^2$)	**168.1**	**4.8**	**140.0**
Client choices and constraints, %			
Sophistication, outright	0.4	33.4	5.1
Counterparties, outright	24.3	0.9	25.5
2*Cov($\widetilde{SPH}_c$, $\widetilde{CP}_c$)	12.2	-0.6	9.3
Market power	36.9	33.7	40.9
OH = Direct trading, outright	15.8	26.2	18.0
2*Cov($\widetilde{OH}_c$,$\widetilde{SPH}_c$)	10.9	29.0	9.5
2*Cov($\widetilde{OH}_c$,$\widetilde{CP}_c$)	28.8	0.3	25.5
Market power+order-handling	**92.5**	**89.2**	**92.9**
Other contributors to variance			
A. Inventories, outright	0.0	0.0	0.0
B. Information, outright	0.3	2.0	0.1
C. Activity discount, outright	0.1	2.0	0.6
D. 2*Covs(costs & benefits)	2.1	2.9	1.9
E. 2*Other covariances	5.0	3.9	3.5
Total, other contributors	7.5	10.8	6.1

Table compares the contributions to cross-client variance in markups according to Equation (2.12):

$$1 = \sum_i \frac{\sigma_{\hat{\imath}}^2}{\sigma_{\widetilde{Mkp}}^2} + \sum_i \sum_{j\neq i} \frac{Cov(\hat{\imath},\hat{\jmath})}{\sigma_{\widetilde{Mkp}}^2}, \qquad i = \{OH, INV, AD, INFM, SPH, CP, SM\}.$$

We calculate fitted markup components for each trade using estimates of Equation (2.11), below, and Equations (2.1’) through (2.10’) in Section (2.6):

$$Markup = \varphi_0+(\varphi_1+\lambda_1)Sz + DT(\varphi_2+\varphi_3 Sz_t)+(\lambda_2+\kappa_2)HL+(\lambda_3+\kappa_3)RSprd+Time(\lambda_4'+\kappa_1')$$
$$+\beta\, CTV + \sum_j InfoCT'_j \pi_j$$
$$+Soph\delta_2'+\theta\, Skew+\psi_{PB}PB + \psi_{ClBk}\, ClBk_{LFD\text{-}NM} + \psi_{SME}SME_{NM}+\psi_{Priv}Priv.$$

For each client we calculate the average of each fitted component. From these we calculate cross-client variances and covariances. Data: all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012.

Table 2.5. Robustness Tests: Sources of Cross-client Variation in OTC Effective Spreads

	30-min. *Info_c*	**Alt. Sophist.**	**Alt. LFD Client Bank**	**Priced by Interbk. Dlr.**
A. All Clients				
Sophistication, outright	4.4	4.4	4.3	24.4
Counterparties, outright	24.2	23.5	23.8	24.5
2*Cov($\widetilde{SPH}_c$,$\widetilde{CP}_c$)	12.2	12.0	12.0	11.4
Market power (narrow)	40.8	35.5	35.8	39.9
OH = Direct trading, outright	15.5	15.2	15.4	19.3
*2*Cov($\widetilde{OH}_c$, $\widetilde{SPH}_c$)*	10.9	10.7	10.7	10.0
*2*Cov($\widetilde{OH}_c$,$\widetilde{CP}_c$)*	28.5	27.8	28.2	13.6
Market power+order-handling	**95.7**	**89.2**	**90.1**	**82.8**
B. Profit-seeking Clients				
Sophistication, outright	35.4	34.0	34.0	21.3
Counterparties, outright	0.9	1.4	1.5	75.9
2*Cov($\widetilde{SPH}_c$,$\widetilde{CP}_c$)	-0.6	-1.1	-0.1	-3.2
Market power (narrow)	35.7	34.3	35.4	42.9
OH = Direct trading, outright	27.1	25.7	26.5	11.5
2*Cov($\widetilde{OH}_c$,$\widetilde{SPH}_c$)	30.5	28.5	29.3	-10.8
2*Cov($\widetilde{OH}_c$,$\widetilde{CP}_c$)	0.3	0.3	0.7	3.9
Market power+order-handling	**93.6**	**88.8**	**91.9**	**85.8**
C. Utilitarian Clients				
Sophistication, outright	5.2	5.3	5.2	26.9
Counterparties, outright	27.7	26.5	27.1	11.0
2*Cov($\widetilde{SPH}_c$,$\widetilde{CP}_c$)	9.7	9.8	9.6	15.7
Market power (narrow)	42.6	41.6	41.9	53.6
OH = Direct trading, outright	18.5	18.5	18.3	23.9
2*Cov($\widetilde{OH}_c$,$\widetilde{SPH}_c$)	9.8	9.8	9.7	11.9
2*Cov($\widetilde{OH}_c$,$\widetilde{CP}_c$)	26.2	25.6	26.0	17.5
Market power+order-handling	**97.1**	**95.5**	**95.9**	**107.2**

Table compares the contributions to cross-client variance in markups according to Equation (2.12):

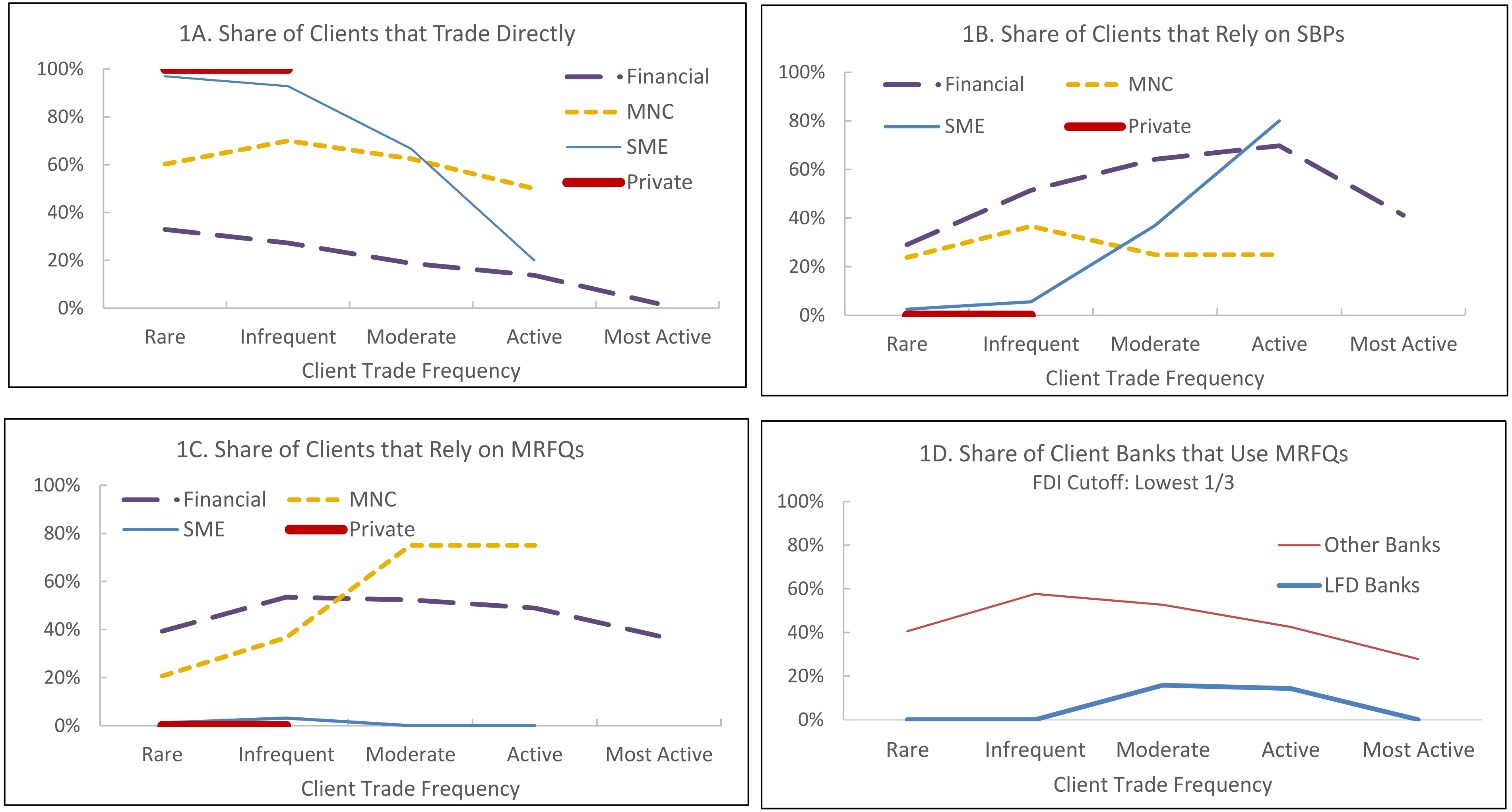
$$1 = \sum_i \frac{\sigma_{\tilde{\imath}}^2}{\sigma_{\widetilde{Mkp}}^2} + \sum_i \sum_{j \neq i} \frac{Cov(\tilde{\imath},\tilde{\jmath})}{\sigma_{\widetilde{Mkp}}^2}, \qquad i = \{OH, INV, AD, INFM, SPH, CP, SM\}.$$

We calculate fitted markup components for each trade using robustness tests for Equation (2.11), below, and Equations (2.1') through (2.10') in Section 2.6:

$$Markup = \varphi_0 + (\varphi_1 + \lambda_1)Sz + DT(\varphi_2 + \varphi_3 Sz_t) + (\lambda_2 + \kappa_2)HL + (\lambda_3 + \kappa_3)RSprd + Time(\lambda_4' + \kappa_1')$$
$$+ \beta\, CTV + \sum_j InfoCT'_j \pi_j$$
$$+ Soph\delta_2' + \theta\, Skew + \psi_{PB} PB + \psi_{ClBk}\, ClBk_{LFD\text{-}NM} + \psi_{SME} SME_{NM} + \psi_{Priv} Priv.$$

For each client we calculate the average of each fitted component. From these we calculate cross-client variances and covariances. Robustness tests are described in notes to Table 3. Data: all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012.

Figure 2.1.. Client Sophistication and Institutional Constraints



Share of clients that rely on a given trading platform, by client trading frequency. *Rare* means no more than one trade per month. *Infrequent* means no more than one trade per week and more than *Rare*. *Moderate* means no more than one trade per day, and more than *Infrequent*. Active means no more than ten trades per day, and more than *Moderate*. *Most Active* means more than ten trades per day. *LFD client banks* come from countries in the lowest half of the IMF's Financial Development Index for 2012. *Other client banks* come from other countries. Data include all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012.

Trading Venue Segmentation and Price Discovery in OTC Markets

3.1 Introduction

An often overlooked but crucial aspect of the foreign exchange market is its reliance on multiple trading venues with varying clearing protocols. These venues differ in speed, transparency, and the informational content of trades, significantly influencing how prices adjust to reflect market conditions. For instance, in request for quote trading venues, dealers know the identity of their counterparties and are able to observe a sequence of their trades (order flow) over time. Such accumulated knowledge enables dealers to learn which trades carry information relevant for the determination of prices (Menkhoff and Schmeling, 2010; Bjønnes et al., 2021). In contrast, trading on order-driven venues allows both banks and customers to submit limit or market orders, with trade information announced on these venues. This transparency fosters a more competitive trading environment and enables quicker integration of information into prices (Hasbrouck, 1991; Bjønnes and Rime, 2005).

This is, to our knowledge, the first paper that analyzes price discovery by disaggregating trading across more than 40 trading venues, addressing a critical gap in the literature. Unlike prior studies, which often concentrated on either the customer or interdealer tier of the market in isolation, this paper explores the dynamics across diverse trading venues in both tiers of the market. Specifically, we examine the contributions of these venues to price discovery, emphasizing the structural aspects in which the venues may differ. Additionally, we utilize counterparty level data to identify which customers have more pronounced effects on prices.

We utilize a unique transaction data set of a large dealer bank covering all transactions executed over a period of more than three months at an intraday frequency. This dataset is particularly well-suited for analyzing price discovery as it captures a wide range of trading activity across diverse venues and customer types, allowing for a detailed examination of interactions in both tiers of the market. The data provide detailed information on every counterparty and include both interdealer and customer trades. Most importantly, we are able to identify four main types of trading venues: a) interdealer order-driven venues, b) multi-dealer order-driven venues, c) request-for-quote venues (interbank, multi-dealer and single-dealer) and d) dark pools.

We use structural vector autoregressions (SVARs) between returns and order flow across different trading venues and customer types, following Hasbrouck (1991). This methodology is particularly well-suited for studying the role of trading venues in price discovery, as it allows for the identification of dynamic relationships between trading venues and price changes while accounting for the timing and directionality of information flow.

We advance the hypothesis that the degree of trading information observable in the market (more commonly known as market transparency) is fundamental for understanding price discovery in the FX market (Bloomfield and O'Hara 1999). Specifically, more transparent trading venues influence the speed at which new information is reflected in prices, and the incorporation of private versus public information into the pricing process. In fact, by using transaction-by-transaction midpoints we find strong evidence that trades on the transparent (order-driven) interdealer trading venues have the strongest price impact. These trades tend to originate from Global Liquidity Providers and Hedge Funds that trade at high-frequency. This information is rapidly incorporated into prices, with the adjustment occurring within five minutes. We also find evidence of much smaller, but still significant, price impact from trades on the order-driven MDPs (multi-dealer trading venues) where dealers trade with customers. This is a finding not sufficiently documented in the FX microstructure literature. Discussions with FX practitioners indicate that they actively monitor a real-time liquidity ladder which aggregates liquidity across dozens of order-driven venues and it is not uncommon for large dealer banks to bypass interbank order-driven venues in favor of MDP order-driven venues.

Except for trades by aggressive financial players, bilateral trades with our bank - such as Request-for-Quote exhibit a negative price impact. This finding may be due to the very low transparency when trading bilaterally with customers which primarily occurs on the dealer's single-bank trading venue (RFQ). We find that the SDP functions as a risk management venue[17], where inventory is initially accumulated rather than immediately offloaded in the market. This suggests that the dealer relies on internalization to manage short-term inventory imbalances before engaging external venues. We hypothesize that risk is reallocated to dark pools or more transparent interbank venues only when exposure levels exceed the dealer's risk tolerance. This aligns with practitioner insights, which suggest that dealers manage inventory according to Value-at-Risk (VaR) constraints.

[17] Menkoff et al. 2016 find evidence of risk-sharing facilitated on large dealer platforms

Finally, we find that price impacts from dark pools are both small and short-lived, suggesting that dark pools primarily facilitate inventory adjustments rather than contributing significantly to price formation. Because anonymity in dark pools conceals the trading intentions of market participants, large dealer banks strategically use dark pools to execute large trades away from order driven platforms, thereby minimizing market impact. Our analysis of Impulse Response Functions across trading venues suggests that some price-relevant information from dark pools may eventually migrate to interbank trading venues, indicating that some limited spillover effects may occur over time.

The paper proceeds as follows. Section 3.2 describes trading venues in the context of the current structure of the FX market, laying the groundwork for understanding the role of trading venues in price discovery. Section 3.3 discusses information in the FX market and the related literature. Section 3.4 presents our data, showcasing its relevance for studying venue-specific dynamics. Section 3.5 presents our empirical results. Finally, Section 3.6 offers some concluding remarks and avenues for future research.

3.2 FX Trading landscape and venues

To better understand the role of trading venues in price discovery, we need some perspective on the clearing protocols that dominate FX trading. Table 6.1.2 (Appendix) categorizes the various types of trading protocols across trading venues and participant tiers. In the customer tier of the market, clients can still contact a dealer directly; however, it has become more common for trades to be executed electronically via the Single Dealer Platform (SDP). The SDP has largely replaced traditional voice sales by transmitting two-way quotes, a process commonly referred to as request-for-quote (RFQ) trading[18]. Because these bilateral customer-dealer trades are only observed by the parties to the trades, dealer banks can condition on the identity of the

18 Our SDP classification contains request-for-quote transactions dealt with the bank's electronic trading venue either a) directly by the customer b) brokered through intermediaries such as electronic brokers and c) executed using an API (Application Programming Interface) connection. An API provides the possibility for customers to connect their systems directly to a dealers' trading station and gain additional functionality, for example receiving a real-time price feed (also called direct executable streaming prices), order management and trade processing services. The API connection has additionally a trading perspective, which allows customers to deploy sophisticated execution options as algorithmic trading strategies.

customer when providing quotes. As a result, transparency and market impact are low compared to other types of FX trading venues.

In contrast to SDP, Multi-Dealer-Platforms that use a request for quote protocol, involve competition among dealers to provide quotes to clients. These trading venues are commonly used by more sophisticated clients such as small banks and large corporate clients, who must satisfy internal requirements regarding best execution which often mandate the review of quotes from multiple dealers before execution.

In order driven trading venues customers can submit limit or market orders and some Multi-Dealer-Platforms allow for customer-to-customer trading (Barclay et al. 2003). Order-driven MDPs provide information about executed trades to participants. In contrast, in bilateral trading on MDPs (such as RFQs) there is no requirement to disseminate trade information to third parties. Interbank venues operate with a central limit order book, which enables pre- trade transparency by providing information about order book depth (size and quotes) and trader identities (such as the quoting counterparty). These trading venues are popular for FX Dealing banks that trade with each other in order to offload residual inventories from customer trading, or to speculate.

Dark pools are an additional venue for dealers where they are likely to prioritize anonymity. Anonymity in dark pools is significant, as trading identities and details of the transaction are concealed until after the transaction. As a result large dealers are able to execute sizeable trades without revealing their intentions to the broader market, thereby minimizing the risk of adverse price movements (Hatheway et al. 2017). Furthermore, restricted access to dark pools ensures that only select participants, often large dealer banks, can interact within these venues, thus reducing the potential for information leakage and adverse selection by informed traders, such as hedge funds. Similar to public crossing networks in equity markets, trades in FX dark pools are matched at a ‘fair mid’ price and they are not publicly displayed (Gomber et al 2016), therefore transparency and market impact remain relatively low. However, execution at these venues is not guaranteed and dealers are likely to use them as complementary to other trading venues.

3.3 Information in FX markets

Information in the FX market is not symmetric, rather it is dispersed across a large number of participants and currencies (Ranaldo and Somogyi 2021). This dispersion means that no single participant holds complete information of the currency's value, and informed traders, such as hedge funds or institutional investors, will often possess private information that allows them to influence price formation. Consequently, price discovery in the FX market relies on the gradual aggregation of this fragmented information through trading activity. In this section we explore how theoretical models of information asymmetry and venue choice provide insights into these dynamics, setting the stage for the empirical analysis that follows.

In the standard literature (Glosten and Milgrom 1985) such information is revealed to the market only gradually through the trading process: uninformed participants will learn about the asset's true value by observing the trades of informed through time. The model assumes that when privately informed investors buy the asset, the price will increase, while the price will decrease when privately informed customers sell the asset.

However, compared with the "real-world" FX market, this model rests on several restrictive assumptions. First, it assumes the presence of a single dealer, whereas the FX market involves multiple dealers operating simultaneously. Second, the model posits that dealers themselves are uninformed, an assumption that conflicts with evidence suggesting that dealers often have access to private information through customer order flows and interdealer trading (Bjonnes et al 2020). These limitations reduce the model's ability to fully capture the complexities and information dynamics of the FX market.

The simultaneous trade-model by Evans and Lyons (2002) captures two important characteristics of the FX market: (i) trading among dealers (interdealer trading), and (ii) private information among dealers. In the FX market, customer order flow is regarded as an important source of private information that accrues to dealers. Furthermore, the model assumes risk aversion, which makes inventory imbalances less acceptable to dealers. Previous empirical studies (e.g. Lyons, 1995, and Bjønnes and Rime, 2005) document that dealers manage inventory quite intensively.

In the simultaneous trade model by Evans and Lyons (2002), dealers quote prices to customers in period one. In round 1, all dealers receive market orders from their own customers. A dealer's customer order is not observed by other dealers, and hence is private information. Interdealer trading takes place in round 2. All dealers quote the same price in round 2 (similar to the round 1 price) since they do not want to reveal their information. In this round, dealers trade among themselves to share inventory risk and to speculate on their private information. At the end of round 2, all dealers observe the net interdealer order flow. The order flows in trading round 2 mirror the customer orders in round 1. In the real world, dealers learn about order flow from brokers too (voice brokers and electronic brokers).

In round 3, dealers use information about net interdealer flow in round 2 to set prices such that the public willingly absorbs all dealer imbalances. The price in round 3 will depend on the total flow that the public needs to absorb, and the public's risk-bearing capacity. If dealers are (on average) long in euro, they must reduce the price for euro to induce customers buy euro. While this model captures well the interactions of dealers and customers in the FX market, it provides limited insights into the crucial aspect of venue choice. Understanding why participants select specific trading venues, whether due to trade size or information asymmetry is critical for understanding price discovery. The literature on venue choice (Barclay et al. 2006, Bessembinder and Venkataraman 2004, Conrad et al. 2003) provides useful insights into venue selection too, by examining trading under different contexts such as order-driven trading venues, RFQ systems, and bilateral OTC trading. In the FX market, dealers may trade directly with other dealers on order driven trading venues, they may trade with other dealers indirectly over Multi-Delaer Platforms or perhaps through bilateral trading with customers (SDP).

Henderschott and Madhavan (2015) study corporate bond markets that share a similar market structure to FX. They provide a framework that attempts to narrow the gap between theoretical models presented earlier and real market dynamics where varying trading protocols coexist. In fact, they study venue choice among traders, in auctions (electronic) dealers compete to provide bids and there is pre-trade and post-trade transparency. In search markets (OTC), dealers interact with traders bilaterally. A key insight is that traders will prefer auction markets for smaller trades where immediate execution will be preferred while search markets are more useful for larger transactions where traders would like to keep information private and limit market impact. By shedding light on the relevance of trading protocols, these findings provide a framework for analyzing how fragmented markets like foreign exchange balance transparency, liquidity provision, and market impact.

3.4 Data

We utilize a dataset containing all spot transactions of a top-10 forex dealing bank over more than three months of trading, from 2 January 2012 to 20 April 2012 (68 trading days). The sample includes over 2 million transactions in 22 currency pairs. Such richness is costly to process, however, and for our purposes a smaller sample will provide ample precision, so we focus on the most liquid currency pair, EURUSD (B.I.S. 2022). A unique feature of the data is that it covers both interdealer trades and end-user trades, and all available trading venues.[19]

The data set differs from data utilized in other studies in several important ways. For instance, some studies use only indicative quotes. However, it is well-documented that transaction prices better reflect market activity because they represented actual transactions unlike indicative quotes, which do not necessarily represent trading intentions (Lyons, 1995). The majority of studies use tradable quotes and transaction data. Below we broadly separate these studies into four categories based on their datasets to provide context of why our data are better suited in studying the dynamics of price discovery . Evans and Lyons (2005), Froot and Ramadorai (2005) use data on aggregated customer trades at daily horizons. Some of these studies (e.g. Evans and Lyons, 2005) have data that separate between some types of customers (e.g. financial vs. non – financial customers). Some other studies use data for a particular group of customers, for example Ramadorai (2008) examines data on institutional fund managers.

Other studies use data from interdealer trading venues. Berger et al. (2008), Moore and Payne (2011), Hau et al. (2002) use interdealer data from EBS or Reuters. Such data lack information on dealer identities.

Other studies in the literature (e.g. Lyons, 1995, Bjønnes and Rime, 2005 and Bjønnes et al., 2021) use data collected from dealers. Such data makes it possible to measure e.g. inventory positions.

Finally more recent papers, for instance Ranaldo and Somogyi (2021) and Ranaldo and Santucci de Magistris (2022) explore data from the CLS which are more representative of the aggregate FX market as they cover a greater number of currencies and counterparties. While these data span longer trading horizons, they lack information on trading venues and the initiating party of the transaction, and therefore order flow cannot be inferred directly.

Our data set is unique since it combines both several customer and interdealer trading venues at high frequency providing a comprehensive view of market activity across different trading venues and counterparties. We are not aware of any other data set in the literature that combines all these features.

For each transaction our dataset provides the following information: Currency pair, date and time stamp of the trade (to the second), transaction price, quantity traded, sign of trade (buy or sell), initiating party, portfolio within the bank to which each trade is assigned, counterparty ID, trading venue and markup relative to the prevailing interbank price.

[19] Internal trades are excluded from the analysis.

Our bank is a major FX market maker so it sets prices in the vast majority trades. Our data set on EURUSD trades contains 210,635 transactions. We partition customers into the following six categories: Global Liquidity Providers (GLP), Regional Liquidity Providers (RLP), Customer Banks, Leveraged Investors (hedged funds and other leveraged investors), Brokers and Corporates clients.

Global Liquidity Providers are banks ranked among the top 20 in Euromoney's bank survey. These banks are active in all major currencies and provide liquidity 24/7 across the globe. Our Regional Liquidity Providers category represents banks that are ranked among the 21-50 in Euromoney's survey. Regional Liquidity Providers are usually a major dealer in their national currency and remain active in EUR/USD trading too.

Our customer bank category is meant to represent banks whose interbank trades are primarily focused on customer service. We define these as banks that are not among the top 50 dealing banks of the 2012 *Euromoney* survey of the foreign exchange market. Customer banks are typically smaller banks from developed economies or medium-to-large financial institutions from emerging economies. They engage in EUR/USD trading mainly to provide liquidity to customers in their small, niche markets and to support their non-dealing operations. Unlike major dealing banks, customer banks do not employ dealers devoted exclusively to trading euro-dollar.

Our leveraged investors category includes not only hedge funds but also commodity trading advisors, algorithmic trading companies, and a few active currency managers.

Our Brokers category includes retail brokers who primarily provide private investors, as well as a limited number of institutional clients, with access to the foreign exchange market. In our Corporates category, clients trade FX to source liquidity for operations and facilitate import and export business. This last group also includes a handful of transactions from private clients and public institutions.

Table 3.1 provides basic descriptive statistics. FX dealing banks represent the majority of trades in our dataset, with Global liquidity providers having the lion share (119,686 transactions). Interestingly, Customer banks have 39,950 trades with our bank, our bank is a main correspondent bank across the globe enabling it to have an extended network of smaller and medium banks. While brokers have 37,436 trades in our data, we note that the average transaction is 507 thousand, which is considerably below the average of ca 1 million for all other client types. We are aware that brokers trade small volumes in high frequency on behalf of retail clients.

Table 3.2 presents trading activity across the main trading venues in our dataset. In interbank trading major dealers trade mainly through order driven venues (35,175 transactions) or reaching out to the counterparty directly through telephone, email etc (2,926 transactions). Single bank platforms are a major trading venue, our dealer provides streaming prices directly on the client's desktop (47,999 transactions). Multi-dealer-platforms enable clients to access liquidity from multiple dealing banks by either a) executing transactions directly from a central order book (order driven MDP) or by using a request-for-quote mechanism that sources near-simultaneous quotes. Dark pools are the most actively traded venue in our dataset (59,813). Large banks trade anonymously on dark pools as a means to offload inventory risk with minimum price impact.

In about 14,000 trades our bank is transacting with Global Liquidity Providers on order driven MDPs. A significant share of these trades are actually executed by customers (typically hedge funds). It is important to note that leveraged investors trade very actively on order driven MDPs and on interbank trading venues, in the case where the leveraged investors are not directly clients of our bank, we observe the name of the prime brokerage bank as the respective counterparty[20].

While leveraged investors in our dataset rely heavily on MDPs, the most significant share of their trading is done through the SDP using an API connection. In Table 3.1 we also present details around the trades with corporate clients. Corporate clients are the least active category of clients in our dataset with ca 1 percent of total transactions. Almost 70 percent of their trades are through the SDP.

The second source of information in our analysis is provided by Reuters and contains quotes and transaction data for the time period covered by the proprietary dataset of our bank. We obtained in particular bid - ask quotes at a transaction frequency, in total 9 737 794 transactions. These data are used for calculating the mid quotes.

3.5 Structural Vector Auto Regressions (SVARs)

To assess whether the exchange-rate impulse from trades is transitory or permanent, we examine how returns respond to order flow using SVARs.

We follow the standard SVAR approach in the literature (Hasbrouck 1991), with applications in FX by Payne (2003), Menkhoff et al. (2010), Bjonnes et al (2021), Ranaldo and Somogyi (2021). For each type of trading venue, we measure order flow as aggressive purchases by the bank's counterparties. Following Hasbrouck (1991), every trade is one observation, so order flow for a specific group is always +1 for an aggressive buy, -1 for an aggressive sale, or 0

20 Prime brokerage arrangements enable highly sophisticated non-dealers such as hedge funds to trade in the interdealer market

(no transaction)[21]. We exclude small trades, that is, trades smaller than EUR 0.25 million. The average and median trade sizes are similar across the three main trading venue types, where as MDP cater for smaller trade sizes (Table 3.2).

The standard model has an underlying structural equation of the form:

$$Ay_t = C(L)y_t + Bu_t \qquad (1)$$

the stochastic error u_t is normally distributed i.e. $u_t \sim N(0,I)$.

and the matrices:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & 1 \end{pmatrix}, \quad y = \begin{pmatrix} dir_t^{IB} \\ dir_t^{DP} \\ dir_t^{MBP} \\ dir_t^{RFQ} \\ r_t \end{pmatrix}$$

$$\text{and } B = \begin{pmatrix} \beta_{11} & 0 & 0 & 0 & 0 \\ 0 & \beta_{22} & 0 & 0 & 0 \\ 0 & 0 & \beta_{33} & 0 & 0 \\ 0 & 0 & 0 & \beta_{44} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

Where dir_t^j represents order flow from the four different venue types and r_t is the log change of the exchange rate mid-price.[22] We use a Cholesky ordering of the form where the midquote is ordered last, hence responding to shocks to all order flow measures. This approach is motivated by microstructure theory as there is little supporting evidence if any for feedback trading (Evans and Lyons 2005, Love and Payne 2008). The groups of order flows are treated symmetrically[23], so they only respond to their own shocks and 8 lags of all variables.[24]

We estimate 6 equations that study the contributions of trading venues to price discovery. Equation (3.1) investigates differences among four broad categories of trading venues: Inter-bank, Multibank, Dark Pools, and Request-for-Quote. Equation (3.2) further disaggregates RFQ flows into: (i) Interbank RFQ, (ii) Multi Bank Platforms RFQ, and (iii) Single Dealer Platforms.

[21] We construct midquotes from transaction prices by adding the average half-spread for sell orders and subtract the average half-spread for sell-orders.

[22] We estimate the model in transaction time that mitigates the clustering of trades which often occurs in high-frequency samples when event time is used.

[23] In each time period there is only one flow and thus no contemporaneous correlation among order flows should exist.

[24] The Schwartz criterion suggests a lag length of 5, while the Hannan-Quinn information criterion suggests a lag length of 8. The results are not sensitive to the exact lag length.

Equations (3.3) through (3.6) examine heterogeneity among trader types. In Equation (3.3), we focus on order flow from banks, distinguishing among Global Liquidity Providers, Regional Liquidity Providers, and Customer Banks. Equation (3.4) explores the interactions between the trading venues specified in Equation (3.2) and banks identified in Equation (3.3). Equation (3.5) isolates the order flow impact from Hedge Funds, and Equation (3.6) integrates order flow from trading venues and both customer groups, banks and hedge funds.

3.6 Results

The results from the SVARs, presented in Table 3.3, highlight the differential impact of order flows across trading venues. Our baseline equations, (3.1) and (3.2), focus exclusively on venue-level dynamics.

In Equation 3.1 (Table 3.3, column 1), we observe that order flows on the order-driven venues (Interbank and MDP OD) exert a significantly stronger impact on midquote returns compared to RFQ and Dark Pool trading venues. In Table 3.3, the estimated coefficients indicate that a one-unit positive shock to trades on interdealer order-driven venues results in a midquote increase of approximately 0.41 basis points, and 0.19 basis points for MDP OD, compared to 0.13 and 0.09 basis points for RFQ and Dark Pool, respectively.

In fact, this information is rapidly incorporated into prices, with the adjustment for Interbank OD occurring within five minutes (see Figure 3.1). The coefficient of 0.41 basis points for Interbank OD venues is economically significant, representing approximately 40% of the average bid-ask spread for trades of similar size[25].

By contrast, a one-unit shock on RFQ trading venues results in a smaller effect of approximately 0.14 basis points. Unlike RFQ, order driven venues offer competitive and continuous trading, and are able to aggregate public and private information more quickly. This finding is consistent with Hendershott and Madhavan (2015), who show that the transparency of order driven trading venues over RFQ is key for their driving role in price discovery.

Results from Equation 3.1 (Table 3.3, column 1) indicate that price impacts from dark pools are both small and short-lived. These findings suggest that dark pools primarily facilitate inventory adjustments rather than contributing significantly to price formation. Because anonymity in dark pools conceals the trading intentions of market participants, large institutional traders seeking low market impact will be more likely to use them (Zhu 2014). This pattern is

[25] Based on market estimates of typical spread levels around 1 basis point

consistent with dealer bank's strategic use of dark pools to execute large trades away from order driven platforms.

In Equation 3.2 (Table 3.3,column 2), we disaggregate RFQ order flow into Single Bank Platforms (SDP) ,Multi-Dealer Platforms (MDP) - RFQ and Interbank RFQ revealing a difference in their effects. While the coefficients on the other venues remain consistent to those in Equation 3.1 (Table 3.3, column 1), the coefficient on SDP order flow is both negative and statistically significant and equal to -0.35. This result suggests that SDP function primarily as liquidity-providing venues rather than platforms for price discovery. The negative coefficient may reflect the nature of order flow that is less informed or directional compared to activity on other trading venues. Transactions executed on SDP typically involve passive participant behavior, reflecting the hedging of foreign exchange exposures or currency conversion for transaction management purposes rather than any speculative or informed trading. These findings align with the broader microstructure literature, which suggests that liquidity driven transactions can lead to price reversal as dealers adjust inventories and manage execution risk (Bjonnes and Rime 2005).

The results of Equation 3.2 (Table 3.3, column 2) confirm that interbank order-driven venues remain central to price discovery with a coefficient of 0.41. The coefficient on order driven MDP with 0.19 exhibits a smaller yet persistent effect on prices and we hypothesize that these venues serve as complementary pricing venues to the main interbank venues. This is a finding not sufficiently documented in the FX microstructure literature. Discussions with FX practitioners indicate that they actively monitor a real-time liquidity ladder, which aggregates liquidity across dozens of order-driven venues, enabling them to identify and execute trades on the venue offering the best execution. In fact, it is not uncommon for large dealer banks to bypass interbank order-driven venues in favor of MDP order-driven venues when seeking superior execution.

3.6.1 Cross Venue Interactions

Our analysis of impulse response functions from Equation (3.2) (Figure 3.2) reveals that, contrary to conventional market microstructure theory, dealers do not immediately offload inventory accumulated from RFQ (SDP) trades into venues with continuous trading, such as Dark Pools or Interbank order-driven venues. Instead, as shown in Panel B, SDP RFQ activity declines following shocks from Dark Pools and Interbank OD, suggesting that SDP RFQ functions as an internalization mechanism, where dealers initially warehouse risk before gradually reallocating it across other venues. In fact, it is now common for more than 80% of major currency flow to be internalized (Chaboud et al. 2023). Further, as illustrated in Panel A, we observe a small but positive impact on Interbank OD following shocks from Dark Pools, indicating that some price-relevant information may eventually migrate from less transparent venues to Interbank venues.

3.6.2 How does price impact vary across client types

While our SVARs using order flow from trading venues provide important insights into price discovery in FX markets, they do not reveal which market participants bring information to the market. Customer order flow may carry information for several reasons. Macro statistics arrive with a time lag, allowing some customers to profitably synthesize existing public information and generate private insights into upcoming developments such as macro statistical releases (Osler 2009). Such customers, particularly hedge funds and other leveraged investors, engage in active research and seek market insights to gain an informational advantage (Bjønnes et al., 2021). The granularity of our data allow us to estimate a number of equations that explore the order flow impacts from specific client groups shedding further light in the role of different market participants in price discovery. Because customer types are segmented across trading venues, we expect the effects to vary accordingly.

In equation 3.4 (Table 3.3, column 4), we expand upon our baseline specification (Equation 3.2) by disaggregating interbank order flow into three categories: global liquidity providers (GLP), regional liquidity providers (RLP) and customer banks which include retail and commercial banks that serve end customers. While our results remain robust across trading venues we observe a strong and persistent price impact from order flow of global liquidity providers. This result is economically significant, the coefficient of 0.47 suggests that order flow from global dealing banks is highly informative - three times larger than that of Customer Banks. This result aligns with Bjønnes and Rime (2005), who demonstrate that such banks are more likely to have access to private information by analyzing customer order flow and operating across a large number of trading venues. In contrast, our results for customer banks with a coefficient 0.15 suggest that their primary role revolves around liquidity provision for smaller end customers, a finding consistent with prior studies (e.g., Mende et al., 2006).

In Equation 3.5 (Table 3.3, column 5), we further disaggregate order flow from hedge funds trading on Multi-Dealer Platforms (MDPs) and Single-Dealer Platforms (SDPs) via API connections. Our findings indicate that hedge fund order flow exerts a stronger and more persistent effect on prices compared to that of other customer groups. Specifically, a one-unit shock originating from hedge funds trading on MDPs results in a permanent midquote adjustment of 0.6 pips whereas the corresponding impact of other customers on MDP is only 0.16 pips.

These findings are also consistent with prior literature from equity markets suggesting that high-frequency trading (HFT) accelerates price discovery. In our dataset, hedge funds utilizing API connections to engage in high frequency trading exhibit an immediate impact on prices. In fact, the effect of their trades is visible within the first minute after the transaction occurs (Figure 3.3). This result aligns with the work of Brogaard et al. (2014), who demonstrate that high-frequency traders contribute to price discovery by interacting with various liquidity providers and trading venues simultaneously.

Finally, in Equation 3.6 (Table 3.3, column 6), we disaggregate order flows from both banks and hedge funds. The results remain consistent with our earlier equations, highlighting that market participants with access to private information exert a permanent effect on prices.

3.7 Conclusion

This paper provides a comprehensive analysis of price discovery in the foreign exchange (FX) market by examining the role of competing trading protocols across multiple trading venues. Using a proprietary dataset of more than 200 thousand EUR/USD transactions from a top 10 dealing bank, we document significant variations in the informational content of trades across interdealer order-driven venues, multi-dealer order-driven venues, RFQ venues, and dark pools.

Our findings highlight that interdealer order-driven trading venues are the most informative, and their transparency facilitates the rapid incorporation of information into prices. In fact, order flow from such venues is incorporated within the first few minutes of trading. While trades on multibank order-driven trading venues also contribute meaningfully to price discovery, their impact is less pronounced. In contrast, RFQ trading venues primarily serve as liquidity provision mechanisms. We also find that the dealer's main single-dealer trading venue (RFQ) acts as a risk warehousing venue, where inventory is initially accumulated rather than immediately offloaded to the market. This suggests that the dealer relies on internalization to manage short-term inventory imbalances before engaging external venues. One such external venue are Dark Pools, where trader anonymity ensures that price impacts will be small and short-lived.

When disaggregating order flows across customer groups, we observe that both hedge fund order flows and those of global liquidity providers exert a stronger and more persistent impact on prices compared to other customer segments. This suggests that these participants may possess superior information for forecasting price movements, potentially due their access to advanced analytics, execution strategies, and the ability to observe aggregate market order flow.

There are several potential extensions to our current study. One promising avenue is investigating how market making varies across different venues. Specifically, this could involve exploring the outcomes of aggressive and passive trading strategies on inventory positions, particularly how these are influenced by the transparency and structure of different venues. Additionally, further empirical work on cross-venue spillovers could provide deeper insights into the transmission of information between trading protocols. Expanding the methodological framework to incorporate alternative structural models, such as network spillover models, could further enhance our understanding of venue interconnectedness, and the broader implications for FX price formation.

Table 3.1: Descriptive statistics per client type

Client Types		Total	GLP	RLP	Hedge Funds	Customer Banks	Brokers	Corporates
Obs.		210,635	119,686	6,545	4,777	39,950	37,436	2,241
Trading Volume	Mean	9.4	9.01	1.37	1.18	1.03	0.513	1.36
(mln)	Max	41.11	81	78	56	184	20	70
	(Stdev.)	-11.24	-6.1	-0.89	-0.57	-1.48	-0.61	-3.98
Venue Shares								
Dark Pools		N	50,685	0	0	0	9,128	0
SDP RFQ		N	1,261	1,516	3,433	27,870	12,320	1,599
Interbank direct		N	138	52	13	2,389	224	110
Interbank O.Driven		N	43,015	2,659	0	3,699	7	38
MDP Order Driven		N	14,010	1,523	400	3,706	9,136	31
MDP RFQ		N	10,577	795	931	2,286	6,621	463

NOTE Table shows summary statistics for different client types in our dataset. Data include all market-making EUR-USD trades by a top-10 EUR-USD dealing bank with dealing-banks and clients during the 68 trading days from 2 January through 20 April, 2012. GLP: Global Liquidity Providers, ranked top 20 in Euromoney survey, RLP: Regional Liquidity Providers, ranked 21-50 in Euromoney survey, Customer banks, smaller retail and commercial banks across the globe either ranked 51-100 in Euromoney survey or not ranked at all, Hedge Funds: Proprietary trading companies employing a combination of strategies: macro, technical and High-Frequency-Trading, Corporates: Aggregate group comprising less sophisticated trading counterparties mostly corporate companies, and a handful private clients and public entities.

Table 3.2: Trade size in EUR millions by venue type

	Total	Interbank OD	Interbank Direct	Dark Pools	MDP Order Driven	MDP Order RFQ	SDP RFQ
Mean	0.97	1.21	1.71	1.00	0.48	0.60	0.97
Median	1.0	1.00	0.55	1.00	0.25	0.28	0.30
Maximum	40	25.00	50.00	5.00	10.00	38.1	184
Std. Dev.	1.69	0.84	0.32	0.95	0.61	0.94	0.32
Observations	195,918	35,175	2,926	59,813	28,402	21,603	47,999

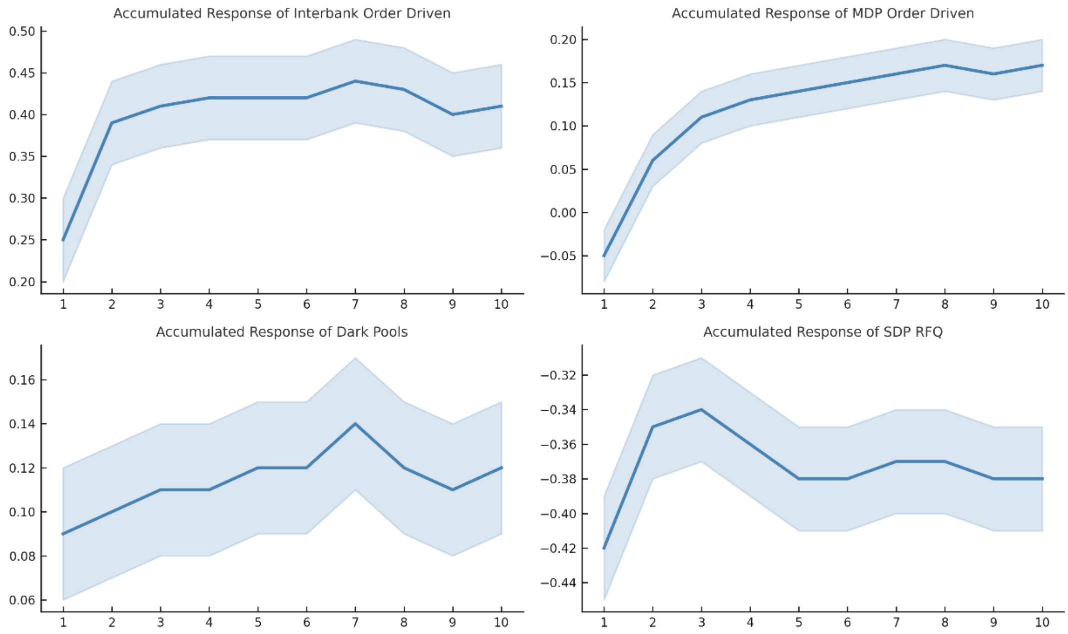
NOTE Table shows summary statistics for trading volume across different venues types in our dataset. Data include all market-making EUR-USD trades by a top-10 EUR-USD dealing bank with dealing-banks and clients during the 68 trading days from 2 January through 20 April, 2012. Interbank order driven: Dealer banks trade with one another on electronic interdealer venues where best bid and ask are visible. Interbank direct: Counterparty reaches out to an interbank dealer by telephone, email, txt, etc. SDP RFQ: Single-dealer platforms provide prices on the client's desk-top and connects directly to the dealing bank. This venue also contains transactions from two additional sources: a) executed on SDP using an Automated Programmable Interface which enables algorithmic trading and b) RFQ bro-kered through an intermediary and traded on SDP. MDP order driven: Multi-Dealer Platforms with a central order book where clients can execute transactions against multiple dealing banks. MDP RFQ: Request for quote systems, one client requests near-simultaneous quotes from multiple dealing banks. Dark Pools, anonymous venues where transactions are executed on mid-quote prices.

Table 3.3 Long-run impulse response

	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)	(3.6)
Interbank OD	0.41	0.41			0.43	
	[0.37, 0.45]	[0.37, 0.46]			[0.41, 0.45]	
GLP			0.48	0.47		0.47
			[0.46, 0.49]	[0.43, 0.52]		[0.45, 0.49]
RLP			0.37	0.36		0.31
			[0.32, 0.45]	[0.21, 0.51]		[0.29, 0.33]
CB			0.15	0.15		0.13
			[0.09, 0.20]	[0.02, 0.27]		[0.11, 0.15]
Interbank RFQ		0.15		0.08	0.11	0.11
		[-0.02, 0.32]		[-0.09, 0.23]	[0.09, 0.13]	[0.09, 0.13]
Dark Pool	0.09	0.09		0.09	0.09	0.09
	[0.04, 0.14]	[0.05, 0.13]		[0.05, 0.14]	[0.07, 0.11]	[0.07, 0.10]
MDP OD	0.19	0.17		0.16	0.16	0.16
	[0.15, 0.23]	[0.11, 0.24]		[0.10, 0.22]	[0.14, 0.18]	[0.14, 0.18]
MDP RFQ		0.12		0.14	0.09	0.09
		[0.05, 0.19]		[0.06, 0.20]	[0.07, 0.11]	[0.07, 0.11]
SDP RFQ		-0.35		-0.34	-0.34	-0.34
		[-0.40, -0.30]		[-0.39, -0.29]	[-0.35, -0.31]	[-0.36, -0.32]
MDP HF					0.60	0.60
					[0.58, 0.62]	[0.59, 0.62]
SDP HF					0.30	0.30
					[0.28, 0.312]	[0.28, 0.32]
All RFQ	0.13					
	[0.09, 0.17]					

NOTE The table reports the long-run impulse responses of midpoint changes to unit innovations in trades, where trades are classified as buy (equal to 1) or sell (equal to -1), alongside their corresponding 95% confidence intervals. The cumulative impulse responses are derived using eight lags. Transactions are restricted to the most liquid trading hours in Europe and the U.S. (approximately 12 hours) and include orders equal to or exceeding EUR 500,000. Variable descriptions as per Table 1 and Table 2. MDP HF captures high-frequency trading activity by hedge funds on multi-dealer platforms. SDP HF tracks high-frequency trading activity by hedge funds on single-dealer platforms. All RFQ consolidates the cumulative impact of RFQ-based trading mechanisms across all trading venues offering RFQ pricing.

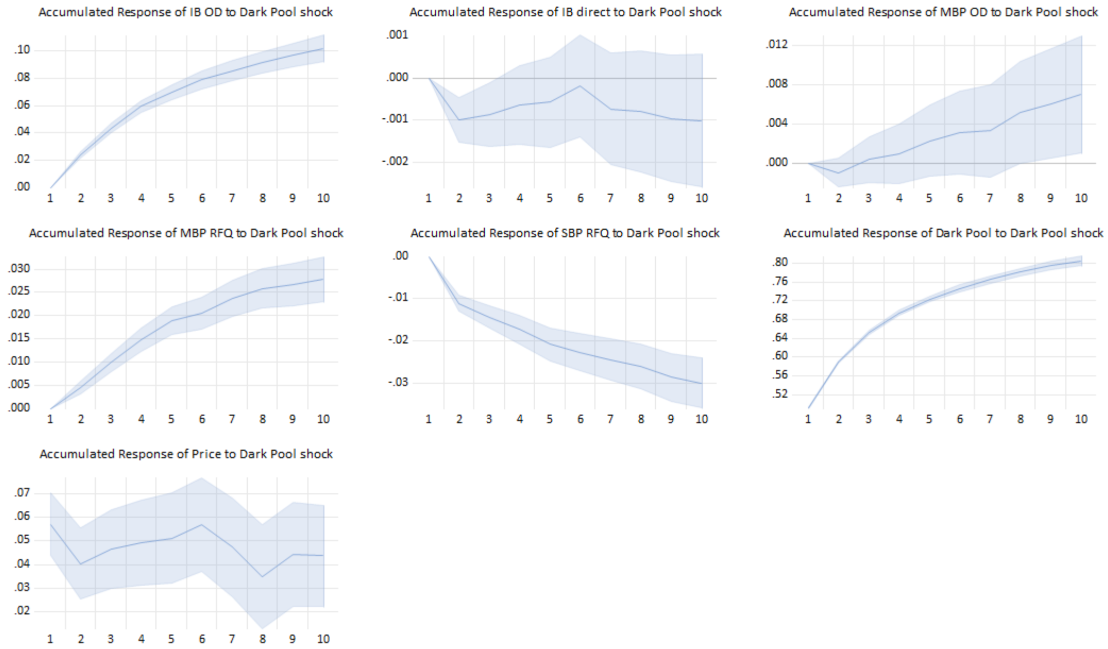
Figure 3.1: Equation 3.2 Impulse Response Functions - Venue Price Impacts



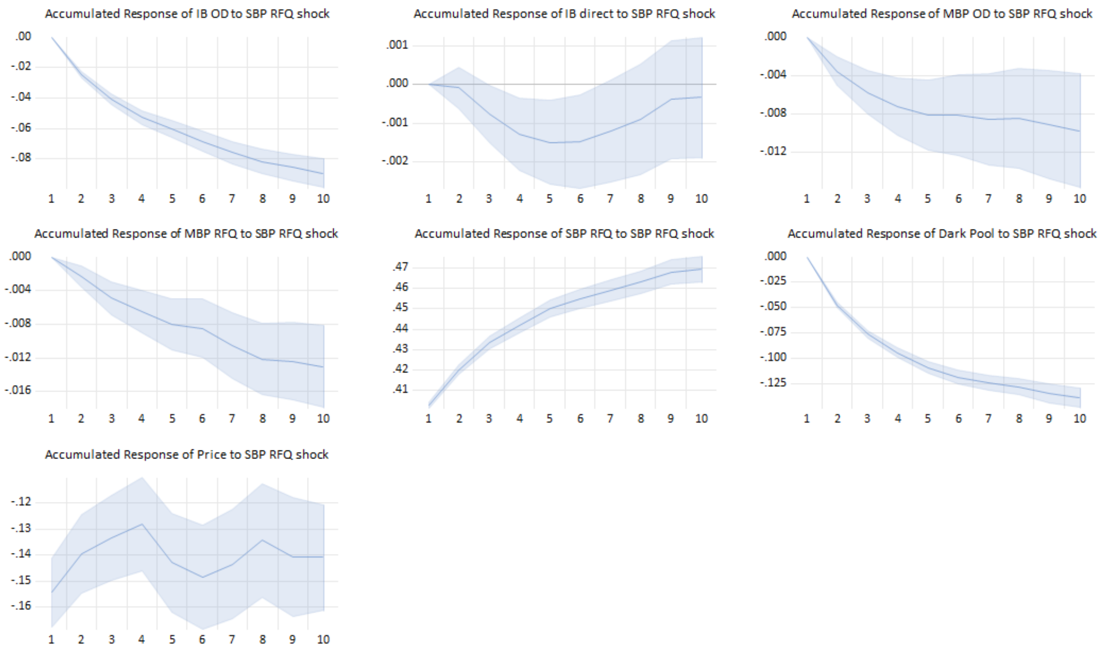
NOTE SVAR analysis of Price Responses to Trades in different trading venues Note: Accumulated impulse response of price to one unit innovations in incoming aggressive trades from counterparties (at 'our' dealers limit orders, Panel A), together with two-standard-deviation bands. (Mid-point) Return (measured in basis points) is ordered last in a VAR together with the five order flows from venues. The five venue order flows are treated symmetrically in that all influence return contemporaneously, but contemporaneous impact on other flows is restricted to zero. Shocks are normalized using the identified structural impact matrix (B matrix) to ensure that impulse responses reflect one-unit structural innovations. Below we present results on the four core venues (MDP RFQ excluded). Vertical axis shows returns in basis points. Horizontal axis is event time (transactions), approximately minutes. Median trade sizes per venue: IB OD = 1m, SDP = 0.3m, MBP OD = 0.25m, Median trade sizes per venue are as follows: IB OD – €1 million, SDP – €0.3 million, MBP OD – €0.25 million, and Dark Pools – €1 million. Equation 3.2 refers to Column 2 in Table 3.3.

Figure 3.2: Equation 3.2 Impulse Response Functions – Cross venue interactions

Panel A

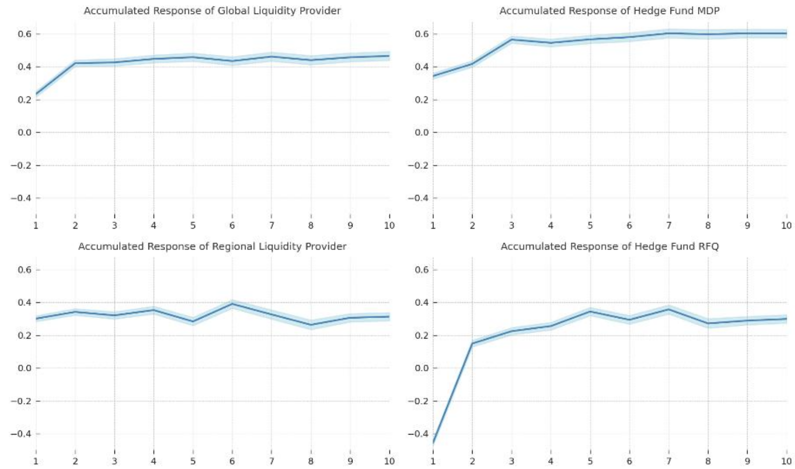


Panel B



NOTE Accumulated impulse response of trading venue activity to one-unit innovations in Dark Pool trading volume (SDP RFQ in Panel B), together with two-standard-deviation confidence bands. Mid-point return (measured in basis points) is ordered last in the VAR alongside the five order flows from venues. The five venue order flows are treated symmetrically in that all influence returns contemporaneously, but their contemporaneous impact on other flows is restricted to zero. Impulse responses reflect reduced-form shocks from the estimated SVAR model without normalization to structural units. The vertical axis shows the cumulative net change in trading activity (i.e. number of additional or fewer trade flow), while the horizontal axis represents event time (transactions), approximately minutes. Equation 3.2 refers to Column 2 in Table 3.3.

Figure 3.3: Equation 3.6 Client Price Impacts



NOTE SVAR analysis of Price Responses to Aggressive Trades by Client type Note: Accumulated impulse response of price to one unit innovations in incoming aggressive trades from counterparties (at 'our' dealers limit orders), together with two-standard-deviation bands. (Mid-point) Return (measured in basis points) is ordered last in a VAR together with the eight order flows. The eight counterparty order flows are treated symmetrically in that all influence return contemporaneously, but contemporaneous impact on other flows is restricted to zero. We present here the IRF from four order flows. Horizontal axis is event time (transactions), approximately minutes. Equation 3.6 refers to Column 6 in Table 3.3.

FRONTRUNNING - DIRECT EVIDENCE FROM A GLOBAL FX LIQUIDITY PROVIDER

4.1 Introduction

We examine how a global foreign exchange liquidity provider adjusts its inventory immediately *before* certain large customer orders are booked. Such trading behaviour by liquidity providers is sometimes labelled ”frontrunning” or ”pre-hedging”. Anticipatory trading strategies are for bidden in many jurisdictions and marketplaces.[126] FX markets are however subject to much less regulation and oversight than many other asset markets, which makes them an ideal setting for a study of anticipatory trading strategies. We use a unique data set from a global foreign exchange dealer bank, and we are able to provide direct evidence on the existence and extent of anticipatory trading practices. This is to our knowledge the first time such evidence is presented. We also provide a novel method to validate the time stamps of financial trade data.

Our analysis consists of two main parts. First, we examine the dealer bank’s inventory adjustment in the time period surrounding trades that are especially susceptible to frontrunning. We find that, for such trades, there is a strongly significant *pre-trade* inventory adjustment, up to 30 minutes before the trade time. We do not find this kind of adjustment for ordinary, non-susceptible trades.

Second, we examine price dynamics before susceptible trades. We find that price dynamics differ significantly before susceptible trades than before other trades. These extraordinary price dynamics are consistent with a price impact stemming from the bank’s observed inventory adjustments.

We illustrate the simple economic intuition of anticipatory trading in a small model. The model shows how profit maximization by the bank leads to anticipatory inventory adjustments and corresponding pattern in pre-trade price dynamics.

[26] For example, in 2018 a US court sentenced the former head of FX cash trading at HSBC to two years in prison for his involvement in the frontrunning of a large currency order placed by an oil company - see e.g. Ex-HSBC FX Trader Sentenced to 2 Years, Sent Directly to Prison, Bloomberg Markets, April 26th 2018.

Figure 4.1 shows our main result on anticipatory inventory adjustment. The figure shows the mean accumulated inventory change of the bank in the 30 minutes before and after the booking times of trades that are particularly susceptible to frontrunning. The ”susceptible trades” in this figure, and throughout our study, are trades with a size larger than 25 million EUR (approximately 30 million USD)[27] that are done either directly between the bank and the customer (e.g. phone, chat, email, fax) or on the bank’s proprietary single bank trading platform (SBP). These are the trades that most likely present a with both an opportunity and the incentive to engage in anticipatory trading, something that we discuss in more detail below. The red bars of the figure shows how the bank’s inventory adjusts *before* the susceptible trade is booked - in the 30 minutes before the bank books a large sale to (purchase from) a customer, we on average see the bank buying (selling) around 12 million EUR. We do not find evidence of a similar pattern for ordinary non-susceptible trades. A statistical analysis of this pattern is performed in Section 4.5.

The rest of this paper is organized as follows. In the remainder of this section we describe and discuss anticipatory trading strategies in more detail, and place the current paper in the context of existing literature. In Section 4.2 we convey the economic intuition of these trading practices, and why they are forbidden in many markets, through a simple sequential trading model. We then describe our unique data set (4.3), our empirical method (4.4) and our results (4.5). The Appendix contains descriptive statistics.

4.1.1 Anticipatory trading

We define anticipatory trading as any trading strategy that relies on certain or uncertain knowledge about other participant’s future trading.

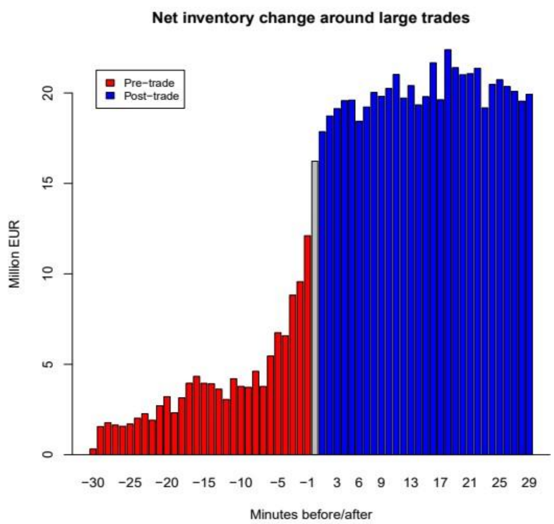
It is important to note that this definition does not necessarily entail a manipulative or other wise illegal trading practice; if for example one market participant is able to infer from publicly available market data that another participant is in the middle of executing a large parent order via trade splitting, it would be rational but not illegal for the detecting participant to engage in anticipatory trading (see Van Kervel and Menkveld (2019) for an example).

[27] We use the 25 million EUR threshold because the bank in question use this number as an internal trigger for triggering human involvement in automated trades. Our results are however robust to the choice of other thresholds.

Anticipatory trading by a dealer would however typically be considered illegal if it involves the use of privileged client information about a future trade to engage in trading intended to benefit the dealer. For example, traders at HSBC were found guilty of fraud in a US court after using the knowledge about an upcoming large customer order to engage in self-serving trading at the expense of their client.[28] Before conducting the large customer order, traders at HSBC took up a proprietary position that benefited from the price impact of the customer trade. This benefit came at the expense of the customer, an oil company, as it got a worse price on its trade than what it would have gotten if HSBC had refrained from frontrunning.

We can not in any certain manner discern between legal and illegal anticipatory trading using our data. As such, we do not make any claims as to the legality of trading patterns uncovered in our analysis. The trading patterns we document can very well be caused by a rational and presumably legal positioning by the bank in anticipation of predictable events, not involving any misuse of confidential customer information.

Figure 4.1 Net Inventory Change around Large Trades



The figure shows accumulated inventory change from 30 minutes before a susceptible trade is booked to 30 minutes after. Inventory change is computed as the cumulative sum of signed traded quantity, excluding all susceptible trades. To reconcile buys and sells in the same plot, we sign trades as follows: when the susceptible trade is a buy, all other sell trades are signed negative and other buy trades are signed positive. When the susceptible trade is a sell, all other sell trades are signed positive and other buy trades are signed negative.

[28] *Former Global Head of HSBC's Foreign Exchange Cash-Trading Found Guilty of Orchestrating Multimillion Dollar Front-Running Scheme*, US Department of Justice, October 23rd 2017.

Nor is there always a clear line separating nefarious trading strategies from the ordinary ones; it is for example common for traders to exchange gossip and rumours about what is going on in the market, and trading in order to position one's inventory based on such rumours may or may not be illegal, depending on the specific circum stances. The distinction between legal and illegal forms of anticipatory trading in the context of US legislation is explored in e.g. Scopino (2014) and Markham (1988).

Anticipatory trading is not a risk-free trading strategy, whether it is based on certain or uncertain knowledge about the future actions of customers or other participants. These trading strategies are based on taking up a position in order to benefit from the price impact of an future trade, but price impact can often be several orders of magnitude lower than price volatility, meaning that any profit generated by an anticipatory strategy can easily be wiped out by non-related movements in the market price.

In our analysis, we focus on anticipatory trading around certain trades, termed "susceptible trades". These are trades of at least 25 million EUR that takes place either via direct bilateral communication with the customer (e.g. chat, phone, fax or email) on through the bank's single bank platform (SBP). We choose to focus on large trades because the incentive for engaging in anticipatory trading is much stronger than before smaller trades; the total price impact generated by a trade is increasing in the size of the trade, and the expected profit generated by anticipatory trading comes from this price impact. Our results are robust to choosing other thresholds than 25 million EUR. We focus on direct and SBP trades because we believe the has more oppor tunity to frontrun these trades: An order that is negotiated over phone clearly provides more opportunity for frontrunning, as does an order that is placed well in advance via email. Since large orders placed on the SBP triggers human interaction and bilateral negotiation, we deem also these trades to be susceptible to frontrunning.

Although there is a large existing literature on frontrunning by intermediaries, there is little in the way of direct evidence. Early literature focused on the behaviour of so-called "dual traders" at the Chicago Mercantile Exchange: Fishman and Longstaff (1992) find evidence of frontrunning, but Chakravarty and Li (2003) use CFTC audit trail data and arrive at the opposite conclusion. Cai (2003) use audit trail transaction data to examine the trading behaviour of market makers in the Treasury bond futures market around the collapse of Long-Term Capital Management in 1998. He finds evidence that market makers engaged in frontrunning, trading for their own accounts 1-2 minutes *before* executing customer orders originating from LTCM.

Comerton-Forde and Tang (2007) find some evidence of frontrunning by a subset of brokers at the Toronto Stock Exchange, but Anand and Subrahmanyam (2008) investigates the same market and concludes there is no evidence in favour of front running. Chaturvedula et al. (2015) finds strong evidence of frontrunning before bulk trades in the Indian stock market.

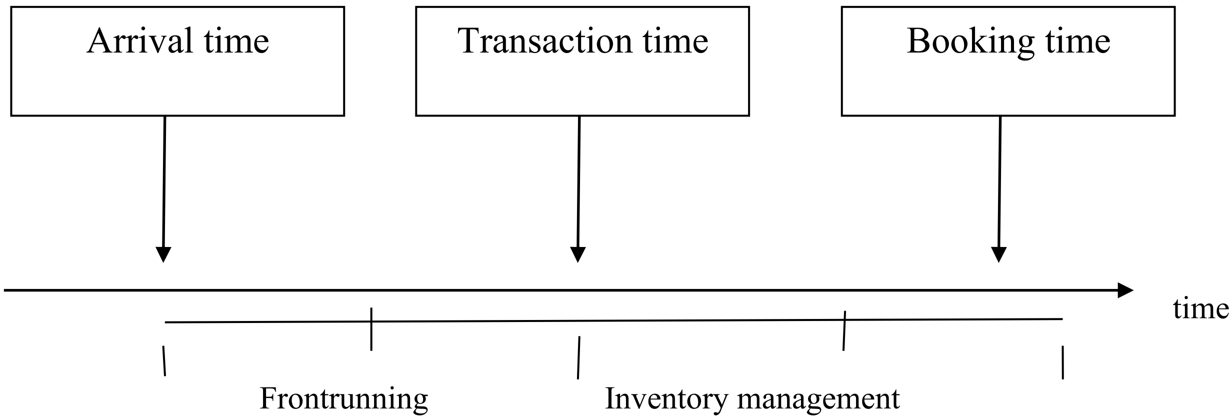
The theoretical literature on frontrunning have focused on welfare implications. Danthine and Moresi (1998) studies frontrunning in an extension of the model from Kyle (1985), and concludes that frontrunning has either no or positive consequences for welfare. Bernhardt and Taub (2008) shows how frontrunning can introduce serial correlation in order flow. Sannikov and Skrzypacz (2014) develops an equilibrium trading model where liquidity providers engage in frontrunning of large investors, and finds that this in general is welfare improving.

4.2 Model

We model a situation where a dealer bank has knowledge of a customer trade before the actual transaction is formalized. This sort of situation arises for example when trades are negotiated over phone or electronic chat. Another common example is customers that follow a fixed or highly predictable trading pattern (for example a customer who buys a large amount of currency every Friday at 1 pm). In our model, the dealer has certain knowledge of the coming trade. In actuality it seems reasonable to assume that the dealer has only a forecast or an informative signal about the size and directions of future customer trades. The focus of our model is solely to convey a simple intuition, and we therefore abstract from this uncertain.

The time line of the model is the following. We refer to the time when the dealer first learns of the coming trade as the *arrival time*. The time when the trade is actually done is the *transaction time*. We refer to the trading activity done by the dealer in the time interval between arrival and transaction as *anticipatory trading* or *frontrunning*, and the trading activity done after the transaction time as *inventory management*. Figure 4.2 illustrates.

Figure 4.2: Timeline of trading activity



We model this situation in an n-period sequential trade-model, inspired by the classic model of Kyle (1985). In each trading period the dealer submits a given order volume that he wishes to trade, and an equilibrium market clearing price results. A related case is analyzed in Saakvitne (2018), in which a model is presented where the market learns about the dealers actions, and price impact is endogenous. Here, we take as given that the equilibrium price is a downwards sloping linear function of the net volume that our dealer wants to transact, meaning that is a price impact from the dealer's trades.

The dealer know beforehand which trading round the large customer trade will occur, as well as the size and direction of the trade. We do not allow the dealer to take up a long-lived speculative position, meaning that our dealer bank starts and ends the model with zero inventory.[29]

The following example illustrates the simple intuition of the model.

Example: Suppose that the initial price of asset A is \$100, and that the model covers three trading rounds. The dealer knows that a customer will sell him 9 units of asset A in trading round two, and that therefore that he will pay his customer the round 2-market price (he does not know what the round 2-market price will actually be). We denote the volume traded by the dealer in the first trading round by x_1, and assume a linear pricing rule with a price impact coefficient of 0.1:

$$p_1 = 100 + .1x_1$$

Similarly, we have $p_2 = p_1 + .1x_2$ and $p_3 = p_2 + .1x_3$.

[29] Osler et al. (2016) analyse a similar model, where the dealer banks is allowed to take up a position of a certain maximum size.

Since the dealer is constrained to end period three with zero inventory, and will buy 9 units from his customer in period 2, he must therefore sell a net amount of 9 units over the course of the three trading rounds.

Consider first the strategy where he does no trading in round one and three, but immediately resell the 9 units the he buys from the customer in trading round 2. The market clearing price in round two will be 100 - .1*9 = $ 99.10. The dealer receives this price from his selling in the interdealer market, and the customer receives this price from the dealer. Thus, the dealer makes no profit.

Consider now instead the strategy where the dealer sells 3 units in trading round one. He receives 100 – .1*3 = $ 99.70 on these units. In round two he sells 6 units, for a price of 99.70 - .1*6 = 99.10. The customer receives $ 99.10 * 9 = $ 891.90 from the dealer, while the dealer earns 99.70*3 + 99.10*6 = $ 893.70. We see that the dealer makes a profit of $ 1.80.

When the dealer makes a profit, logic requires that someone else must have made a loss. It is the customer who loses whatever the dealer profits, because the customer receives a worse price on his trade than what he would have done if the dealer did not frontrun him. Here we see the rationale for why frontrunning is forbidden in many markets - the dealer is profiting from his information about his customer's future trading interest, at the expense of the customer. In certain cases this would be deemed a misuse of inside information by market regulators.

Let us now take the perspective of the dealer, and ask whether we can do even better. It turns out that there exists a unique optimal trading strategy, which involves selling *more* than the total customer order in periods one and two, and buying back the remainder in period three: the optimal thing to do is to sell 6 units in period 1, sell 6 units in period 2, and buy 3 units in period three.

We now derive a formal model, building on the intuition of the previous example. Suppose the price in period t can be written as a linear function of the previous price, the volume traded and a white noise-term ϵ_t ,

$$p_t = p_{t-1} + \lambda x_t + \epsilon_t \qquad (4.1)$$

There are a total for N trading rounds, and the dealer knows that in period k he will buy y units from a customer. The model extends to the case where the customer is buying by allowing y to take negative values.

Denote the inventory of the dealer by X_t. The dealer's trades are denoted (x_n, $n = 1, \ldots,$ N). Since the dealer is not allowed to take up a long-lived speculative position, we have $x_0 = 0$ and the constraint

$$\sum_{n=1}^{N} x_n = y \tag{4.2}$$

constraintX^N n=1

The profit of the dealer is simply the income he earns on his own trades in the interdealer market, less the price he pays to his customer:

$$\pi\,(x_1, \ldots, x_{N=} \sum_{n=1}^{N} p_n x_n - p_k y \tag{4.3}$$

The dealer in our model maximizes a mean-variance type objective function:

$$U(\pi) = \mathrm{E}[\pi \mid y] - \gamma \mathrm{var}(\pi \mid y) \tag{4.4}$$

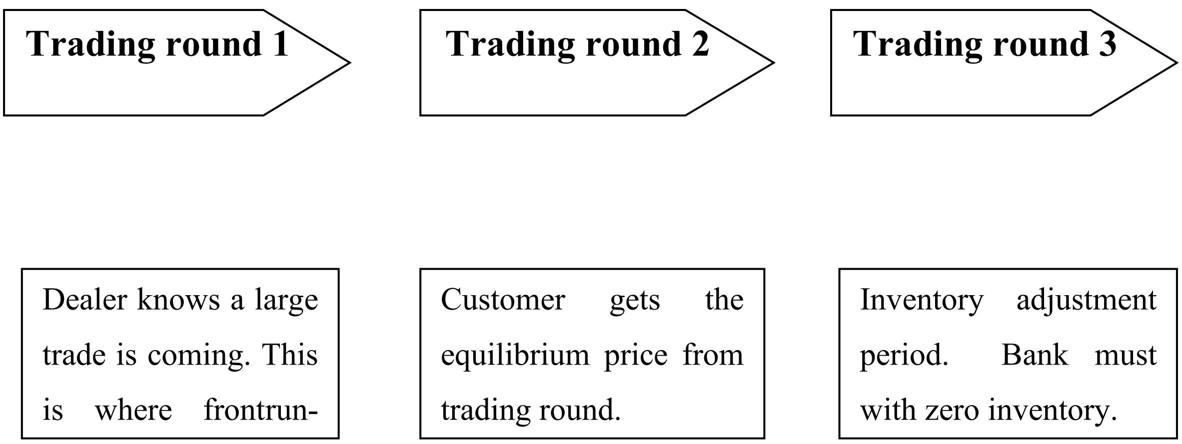
The general model is solved by numerical methods, but the simple case with a risk neutral dealer and three trading periods ($N = 3$, $k = 2$, $\gamma = 0$) is straightforward to solve analytically, and we do so in the next section.

4.2.1 Risk neutral dealer in a three-period model

When the dealer is risk neutral ($\gamma = 0$), he maximizes expected profit. When in addition there are only three trading periods, one before and one after the transaction time, the model is straightforward to solve: one inserts the pricing function and the inventory assumption into the expected profit function, and solves the first order condition of a maximum:

$$\frac{\partial \mathrm{E}[\pi]}{\partial x_1} = 0, \frac{\partial \mathrm{E}[\pi]}{\partial x_2} = 0$$

Figure 4.3: Timeline of the three-period model



Solving the resulting system of two equations yields

$$x_1^{*}= -\ \frac{2}{3}y,\ x_2^{*}= -\ \frac{2}{3}y,\ x_3^{*}= +\ \frac{1}{3}y$$

This optimal strategy has three notable features, the first two of which also carries over to the general model.

First, it is optimal for the dealer to start his trading before the large customer transaction is finalized (i.e. before the transaction time). One way to interpret this is that the dealer becomes informed once he learns about the coming customer order - because there is a price impact in this market, the dealer has a non-zero expectation about future short-term price changes. Trading on such information is the very definition of frontrunning, as discussed in Section 4.1.

Second, we see that the dealer is *overtrading*, meaning that he trades more than the total customer order in the two first periods taken together. He then offloads his remaining inventory in the final trading period.

Third, the optimal trading strategy of the dealer does not depend on the price impact parameter λ. It is however straightforward to verify that his expected profit increase linearly in λ. It is the existence of a price impact from trading that enables the frontrunning strategy. It follows that the dealer has stronger incentives to frontrun when the price impact is high. For the same reason the dealer would ideally like that the price impact is high when he has the opportunity to frontrun his customer.

4.2.2 The general model

We solve the general *n*-period model with risk aversion by applying a sample average approxi mation to attain the dealer's optimal solution. See Homem-de Mello and Bayraksan (2014) for a review of this method and its theoretical properties.[30]

[30] Semi-closed form solutions to a related model is given in Osler et al. (2016). If one is willing to use a continuous-time formulation, the model can be made analytically tractable by an application of Ito's formula, and one can then proceed to solve via deterministic optimal control theory - please contact the author for more details on this.

Figure 4.4 show model-predicted inventory for the dealer in the trading periods before and after the transaction time *k*, for a unit customer order. The effect of the customer order itself is not included in these figures, in order to make them comparable to the corresponding empirical plot in Figure 4.1.

As shown in Figure 4.4a, when the dealer is risk neutral and has 30 trading periods, he uses the trading periods before the transaction time to gradually build up an inventory of seven times the size of the actual customer order. His inventory reaches a maximum at the transaction time, and then linearly declines until he satisfy the zero-net position constraint when the model ends.

Figures 4.4b-4.4d show the three effects of risk aversion:

First, the amount of overtrading done by the dealer declines rapidly in the degree of risk aver sion. When $\gamma = 0.01$ the dealer builds up a total inventory around twice the size of the customer order, while for $\gamma = 0.1$ and $\gamma = 1$ the total position is just barely larger than the customer order.

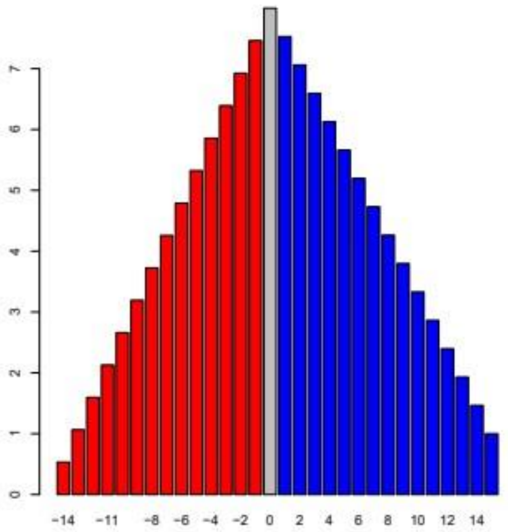
Secondly, the rate of inventory accumulation before the transaction time is also rapidly decreas ing in the degree risk of risk aversion. When risk aversion is high ($\gamma = 1$), the dealer engages in just a tiny amount of frontrunning, and only in the trading period right before the transaction time.

The third effect of risk aversion relates to the unwinding of the dealers proprietary position, meaning how he reduces his position back to neutral and satisfies his zero net-position constraint. The larger the risk aversion coefficient, the faster is the unwinding.

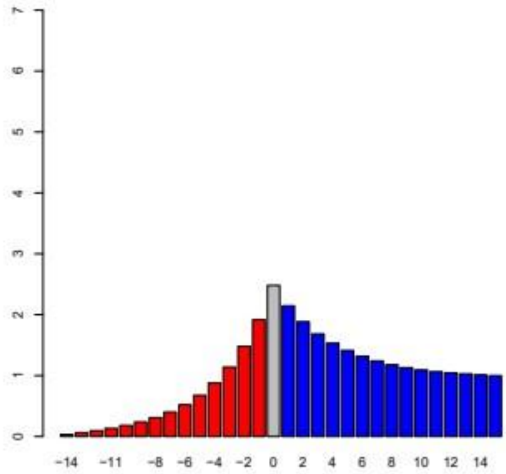
Figure 4.4: Model-predicted inventory change for various degrees of risk aversion.

N=30, k = 15, $\sim N(0, 1)$, *y* = 1.

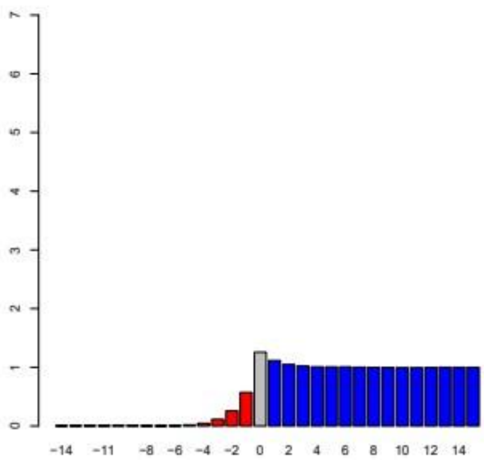
(a) $\gamma = 0$



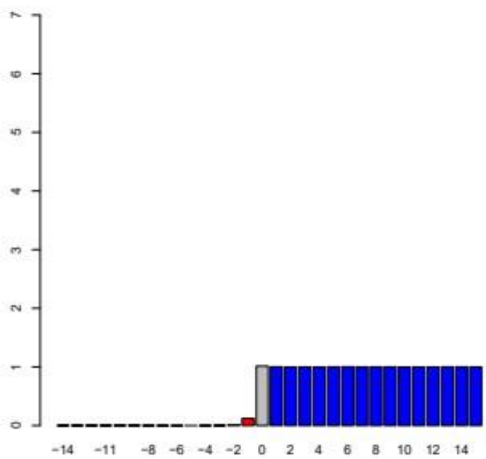
(b) $\gamma = 0.01$



(c) $\gamma = 0.1$



(d) $\gamma = 1$



4.3 Data

Our data come from the largest OTC market, spot FX, where trading averages USD 1.7 trillion per day (Bank of International Settlements, 2016). The dataset contains all spot deals of a top-10 forex dealer bank in EUR/USD (Euromoney FX survey, 2012) over more than three months of trading, from 2 January 2012 to 20 April 2012 (68 trading days). The sample includes

471,844 transactions. EUR/USD is the most liquid currency pair with a daily turnover of roughly USD 400 billion (Bank of International Settlements, 2016).

Our data set is unique since it combines several customer and interdealer trading venues at high frequency. We are not aware of any other data set in the literature that combines all these features. For each transaction, the data provides the following information: currency pair, date and time stamp of the trade (to the second), transaction price, quantity traded, sign of trade (buy or sell), initiating party, portfolio within the bank to which each trade is assigned, counter party ID, trading venue (e.g. Hotspot), prime broker (if any), and markup. Customer markups are measured in pips ($0.0001 per euro) relative to the prevailing core price. The core price reflects information on price and depth from several trading platforms such as EBS, Reuters and Hotspot. Our bank is a major FX market maker so it sets prices in the vast majority trades. The data contains trades with all types of customers such as retail brokers, hedge funds, real asset managers, corporates, and small banks regarded as customers (ranked below top 50 by Euromoney FX Survey, 2012). In addition, the data also includes all trades with FX trading banks (ranked top 50).

These data are well suited for examining frontrunning. Most data sets focus exclusively on dealer-to-customer trades (D2C) or dealer-to-dealer trades (D2D), and provide much less details than our dataset. For instance, Evans and Lyons (2002) and Bjonnes and Rime (2005) focus only on the D2D trades, while Evans and Lyons (2005) and Froot and Ramadorai (2005) focus on D2C trades. The OTC datasets that do include D2C trades often provide no information about customers (Green et al., 2007), or they group customers into a few broad types, such as financial and commercial (Evans and Lyons, 2005). Our data, by contrast, provides individual customer identifiers and we can divide customers into several customer categories. With data covering all trades, including both the customer and interdealer market, we can also track inventory changes. In order to study frontrunning, this is an important feature of the dataset.

The data are also well-suited for analysing today's FX market since they identify over 20 trading venues; most other FX dataset are limited to interdealer platforms (Evans and Lyons, 2002) or they predate the market's fragmentation (Osler et al., 2011). Our empirical analysis is carried out on a trade-by-trade basis and customer-by-customer basis and fully exploits the customer identities in our data. Other features of the same data set are explored in Bjonnes et al. (2017).

Further descriptive statistics on the main data set is reported in Section 4.7 of the Appendix.

The second source of information in our analysis is provided by Reuters and contains quotes, depth and transaction data for the time period covered by the proprietary dataset of our bank (frequency is 1/10 sec.). We calculate midquotes (and depth) using the data from Reuters in order to estimate trading costs at different times.

4.4 Methodology

We report two main pieces of analysis in this paper. The analysis examines how the dealer bank adjusts its inventory before and after certain customer trades, with a particular emphasis on the anticipatory adjustment. The second analysis examines price dynamics before the same trades. In addition, in Section 4.6 we describe and report a third piece of analysis, aimed at veri fying the validity of the time stamps in our data. This analysis can be seen as a robustness check.

The model of Section 4.2 has predictions on both pre-trade inventory adjustments and pre-trade price dynamics. If the bank is engaging in anticipatory trading, we would expect to see a significant build-up of inventory before certain susceptible trades are booked. As a result of this build-up we would also expect to see, for these particular trades, market prices being driven down before a customer sells, and up before a customer buys.

4.4.1 Inventory analysis

The aim of the inventory analysis is to examine changes to the banks inventory position after, and more importantly, *before* the transaction times a trade susceptible to frontrunning is booked.

In Section 4.1 we discussed which characteristics of a trade that gives a liquidity provider opportu nity and incentives to engage in anticipatory trading. For the purpose of the inventory analysis we operationalize these characteristics as the set of trades satisfying the following two criteria:

1. Trade size of 25 million Euro or more
2. Trade occurs either on the bank's single bank platform (SBP) or via direct bilateral communications with the customer (e.g. phone, electronic chat, email etc.)

The set of trades satisfying these criteria are referred to as ”susceptible trades”. The inventory analysis is performed as follows. Let $f(t)$ be the net inventory change of the bank over the time period $(t, t + 1)$, where t is a time stamp given in whole minutes (for example 2012-01-01 14:15:00) . We denote the pre-trade inventory change at time lead x as $y_x(t)$,

$$y_x(t) = \begin{cases} (+1)\sum_{n=1}^{x} f(t-n) & if\ customer\ buys \\ (-1)\sum_{n=1}^{x} f(t-n) & if\ gustomer\ sells \end{cases} \tag{4.5}$$

Defined in this way, $y_0(t)$ is the inventory change over the minute where susceptible trade took place, $y_5(t)$ is the inventory change over the 5 minutes before the trade took place, and so on. The sign change for trades where the customer sold to the bank ensures that we can treat buys and sells in the same manner later in our analysis.

Let $T = (t_1, t_2, \ldots, t_N)$ be the trade times for the N susceptible trades, rounded down to the nearest whole minute. If the bank is adjusting its inventory in anticipation of large customer trades, we expect $y_x(t)$ to be positive for the susceptible trade times $t \in T$. In Table 4.1, we report the mean $\bar{y}_x$ for all susceptible trades, meaning

$$\bar{y}_x = N^{-1} \sum_{t=t_1}^{t_N} y_x(t) \tag{4.6}$$

To test the statistical significance of average pre-trade inventory changes $\bar{y}_x$, we seek to compare the inventory change for susceptible trades with the corresponding change for ordinary trades.

To this end, we compute a bootstrap distribution of average pre-trade inventory changes for *any* trade, not only the susceptible trades.

The bootstrap distribution is built as follows. Let $H = (h_1, h_2, \ldots, h_K)$ be the time stamps of all trades. Table 4.4 shows there are $K = 275, 251$ trades in our sample. One run of the bootstrap procedure involves drawing $N = 74$ trade times at random from H, and computing $\bar{y}_x$ per equations (4.5) and (4.6) for these trades. We perform a large number of such runs to build up a bootstrap distribution of $\bar{y}_x$ under the null hypothesis that pre-inventory changes for susceptible trades are no different than from any other trade. A bootstrap p-value can then be found by comparing $\bar{y}_x$ computed on the susceptible trade set against the quantiles of the distribution $F_{\bar{y}_x}$.

4.4.2 Pre-trade price dynamics

If the bank is engaging in frontrunning, we expect there to be a pattern to market prices in the period preceding especially large trades, caused by the price impact of the bank trading in anticipation of the large trade. We investigate the presence of such a pattern by measuring changes to the midpoint price in the Reuters reference data.

Let $m(t)$ be the reference midpoint price at time t. For a given trade occurring at time t, we define the pre-trade price impact at time lead x, $z_x(t)$ as

$$z_x(t) = \begin{cases} m(t) - m(t-x) \ if\ customer\ buys \\ (m(t-x) - m(t)\ if\ customer\ sells) \end{cases} \quad (4.5)$$

In our empirical investigation we consider time leads x of 0, 1, 5, 10, 15, 20, 25 and 30 minutes.

The pre-trade price impact z is used as dependent variable in a regression analysis. As independent variables of interest, we consider trade volume (unsigned), and a dummy variable indicating whether the customer is a broker/bank or not[31] In addition we use daily volatility as a control variable. Daily volatility has been standardized for interpretability. For a given time lag x we estimate the regression equation

$$z_x(t) = \alpha + \beta_1 \text{VOLUME}(t) + +\beta_2 \text{VOLUME}^2(t) + \beta_3 D_{NONBANK}(t) + \beta_4 \text{VOL} + e_t \quad (4.7)$$

We estimate these regressions using both ordinary least squares and, as a robustness check, using quantile regressions. The regressions are estimated time lags (x) of 5, 10 and 30 minutes.

We expect frontrunning to be more pronounced for large trades, since these trades provide a stronger incentive for anticipatory trading. We also expect frontrunning to be more pronounced for trades with clients that have a more predictable trading pattern, such as corporation and funds. We shall therefore test whether the coefficients β_1 and β_3 are significantly different from zero, and interpret rejection of the null hypothesis as evidence for frontrunning.

[31] The non-banks in this sample would be real money investors, hedge funds and large corporations

4.5 Results

Both the inventory analysis and price change analysis provide evidence consistent with frontrunning, and in particular the model in Section 4.2.

4.5.1 Pre-trade inventory effects

There is a pre-trade inventory adjustment before large trades involving human interaction that is significant at the 1% level, for time leads of 10, 15 and 30 minutes. This is seen in Table 4.1. The red bars of Figure 4.1 illustrates the build-up of this inventory adjustment over the 30 minutes before large trades.

The pre-trade inventory adjustments seen from Table 4.1 and Figure 4.1 are consistent with both economic intuition and the frontrunning model presented in Section 4.2. We do however not find robust evidence for the overtrading and subsequent inventory reversal predicted by the model. The lack of such evidence can be seen visually in Figure 4.1.

Table 4.1: Mean pre-trade inventory changes and quantiles

	$x = 5$	$x = 10$	$x = 15$	$x = 30$
1%	−1.78	−2.02	−2.06	−2.72
5%	−1.12	−1.37	−1.49	−2.15
50%	−0.04	0.16	0.47	0.87
95%	1.42	1.86	2.62	4.96
99%	1.87	2.42	2.65	6.07
$\bar{y}_x$	0.46	2.73	5.25	7.79
	0.46	2.73	5.25	7.79

There can be several reasons for the lack of evidence in favour of overtrading and subsequent inventory reversal. One possibility is that the bank does not engage in this aspect of frontrunning, perhaps because the strategy is deemed as involving unacceptable levels of reputational risk, or as being non-compliant with internal or external regulations. There would also be market risk associated with the overtrading predicted by our simple model - the bank would be exposed to adverse movements in the spot rate above and beyond what is implied by ordinary market making. The theoretical model in fact shows (Figure 4.4) that the extent of overtrading decrease in aversion to market risk.

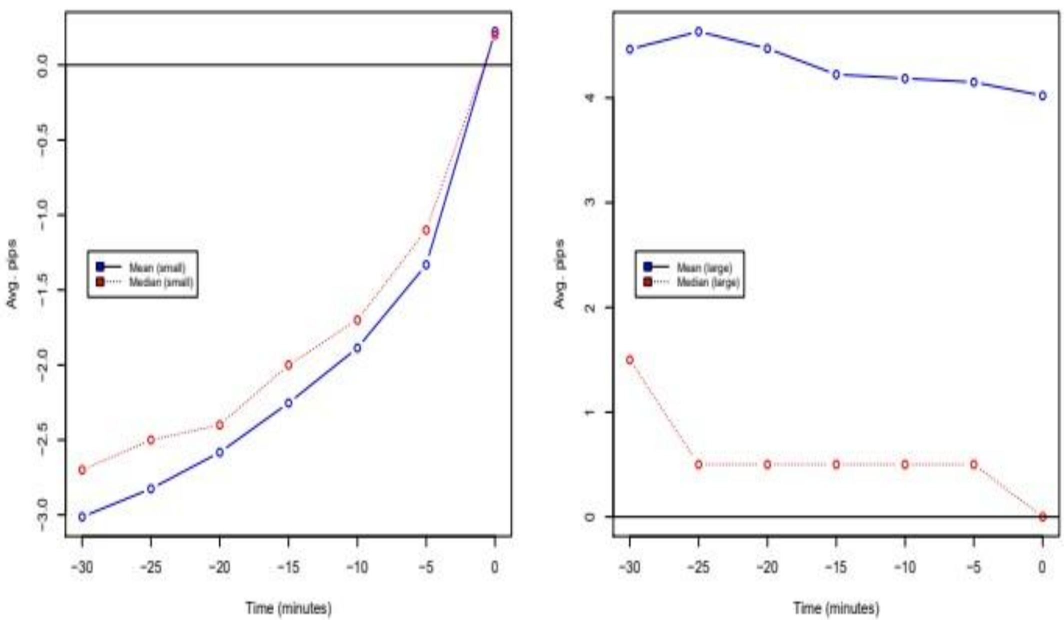
Another possibility for why we are not seeing overtrading is that significant parts of the bank's inventory adjustment is done through either other currency pairs or through internal netting. There are facts supportive of this interpretation: the mean size of the susceptible trades in our sample is 72 million Euro, while the mean *post-trade* inventory adjustment hovers around 8 to 10 million Euro for various time lags (shown in the blue bars of Figure 4.1).

4.5.2 Pre-trade price effects

Figure 4.5 shows average and median trading costs in two samples - small trades (4.5a) and large trades (4.5b) - and illustrates our results on pre-trade price effects. Small and large trades are those with trade size below and above 25 million EUR respectively.

The figure shows two interesting things. First, we see that the market midpoint typically changes in the customers disfavor in the 30 minutes preceding a large trade. If the average customer is benchmarking his trading costs against the prevailing market midpoint at the time of the trade ($x = 0$), his average trading cost would be a little less than 6 pips. If, on the other hand, the customer benchmarks his trading costs against the prevailing market midpoint 20 minutes before the actual trade ($x = 20$), he would find his trading costs to be a little over 8 pips. The difference of 2 pips can be defined as the frontrunning costs.

Figure 4.5: Customer spread versus market midprice

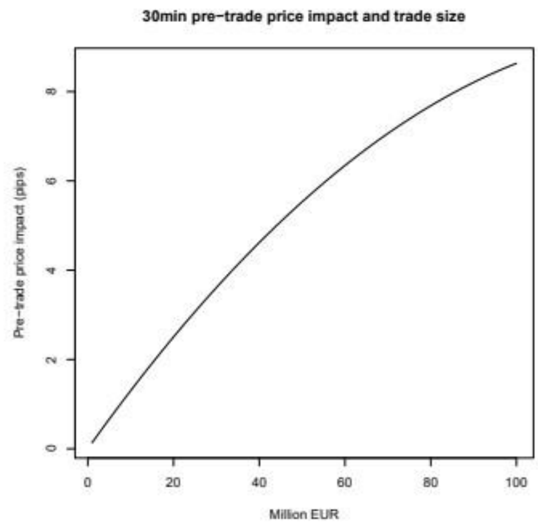


(a) Small trades (b) Large trades

The second interesting thing that can be seen from figure 4.5 is that the average customer seems to be timing his trade in a particular manner, evident from the upwards slope of costs for the full sample of trades. The average buy trade is placed after a period of falling market (mid)price, the average sell trade is placed after a period of raising market (mid)price. This pattern is intuitive – when the market price goes up it triggers profitable trading opportunities for market participants waiting to sell (for example related to hedging motives, information, speculation and so on), while when the market price goes down it triggers profitable trading opportunities for market participants waiting to buy. This pattern is interesting because it implies that for the average trade, pre-trade price impact is negative, which makes the positive pre-trade price impact for large trades all the more striking.

Table 4.2 summarizes the results from estimating the regression equation (4.7). We see that there is a significant size effect to pre-trade price impact, of approximately 0.15 pips per million EUR. The coefficient on squared volume means that the relation between size and pre-trade price impact is concave; the relation is plotted in Figure 4.6. The positive coefficient on the dummy ”Non-liq.provider” indicates that there is significantly higher pre-trade price impact when the customer is not a broker or a bank, which in our data means that it is classified as a real money investor, a hedge fund or a multinational corporation.

Figure 4.6: Estimated relation between trade size and pre-trade price impact



Note: Price impact measured in basis points. 30min pre−trade price impact and trade size

Table 4.2: Pre-trade price impact regression models –

	30 minutes	10 minutes	5 minutes
Intercept	-3.4458^{***}	-2.3224^{***}	-1.7427^{***}
	(0.0621)	(0.0388)	(0.0283)
Size	1.1474^{***}	0.1814^{***}	0.1728^{***}
	(0.0264)	(0.0165)	(0.0120)
Size squared	-0.0006^{*}	-0.0011^{***}	-0.0011^{***}
	(0.0003)	(0.0002)	(0.2347)
Non-liq.provider	2.0084^{***}	-1.6117^{***}	-0.7880^{***}
	(0.5163)	(0.3227)	(0.2278)
Hourly vola.	-2.2424^{***}	-1.2644^{***}	-0.7880^{***}
	(0.0612)	(0.0382)	(0.2278)
Num. obs.	84852	84849	84854

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Note: OLS regression, . Units measures in pips.

4.6 Robustness - are the timestamps reliable?

4.6.1 Time lag analysis - A method for validating time stamps in trade data

The time lag analysis is a methodology developed to ensure that the time stamps in our data are correct; and in particular that large trades are not booked with a time lag. This methodology can also be applied in other contexts than the current paper, so we find it useful to formulate our problem in more general terms.

Table 4.3: Pre-tradeprice impact regression models

	30 minutes	10 minutes	5 minutes
Intercept	-4.3901^{***}	-2.5215^{***}	-1.6797^{**}
	(0.0906)	(0.0486)	(0.0306)
Size	0.1170^{**}	0.1746^{***}	0.1553^{**}
	(0.0386)	(0.0199)	(0.0118)
Size squared	-0.0003	-0.0001^{**}	-0.9848^{*}
Non-liq.provider	2.2420^{**}	1.7520^{***}	0.9848^{**}
	(0.5318)	(0.2599)	(0.2139)
Hourly vola.	-3.0928^{**}	-1.4405^{***}	-0.8053^{**}
	(0.0936)	(0.0557)	(0.0356)
Num. obs.	84852	84849	84854

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

.

Note: Quantile regression (median) Units measured in pips

The general problem looks as follows. Suppose we have records of *N* financial trades of a particular currency pair. Sells and buys are distributed evenly amongst the recorded trades. Each record *i* consists of a transaction price v_i and a time stamp t_i. Each trade is timestamped when it is entered into the electronic bookkeeping system, a process called "booking" the trade. How ever, the researcher suspects that some time goes by from when a trade is actually done (the transaction time of Figure 4.2) until it is booked. In other words, the trades might be booked with a time lag.

Suppose further that we have available a time series $m(t)$ of market mid point prices of the currency pair in question. The transactions in the trade records are thought to take place at the market mid price plus/minus a markup. The markup may vary from trade to trade, but is believed to be zero on average, as there is a roughly equal number of buys and sells in the transaction records.

We model this situation by assuming that each trade *i* actually took place at time $t_i - \theta$, where θ is a non-negative parameter to be estimated from the data. We also assume that each transaction takes place at the prevailing market price plus a Gaussian white noise component (the markup):

$$v_i = m(t_i - \theta) + \epsilon_i$$

$$\epsilon \sim N(0, \sigma_\epsilon^2)$$

Further, we model the market mid price of the currency pair $m(t)$ by a Brownian motion $B(t)$:

$$m(t) = \sigma B(t)$$

To estimate θ, define for each *i* the pricing error $\Delta_i(x)$:

$$\Delta_i(x) = v_i - m(t_i - x)$$

$$= m(t_i - \theta) + \epsilon - m(t_i - x)$$

$$= \sigma [B(t_i - \theta) - B(t_i - x)] + \epsilon$$

$$\sim N(0, \sigma^2|\theta - x| + \sigma_\epsilon^2)$$

It is clear that E[Δ_i (x)] = 0 for all x. It is also clear that the variance of Δ_i (x) is at a minimum when $x = \theta$, so that

$$\theta = \underset{x}{agr\,\min}\, \mathrm{E}[(\Delta_i(x))^2]$$

We approximate the expectation by the sample mean, which gives the estimator $\hat{\theta}$:

$$\hat{\theta} := \underset{x}{agr\,min}\, \frac{1}{N}\sum_{i=1}^{N}\Delta_i(x)^2 \tag{4.8}$$

The estimator $\hat{\theta}$ is an extremum estimator, a class of estimators which are known to be asymp totically consistent also under much weaker conditions than what we have assumed here (Newey and McFadden, 1994).

Figure 4.7a maps a simulation of the root mean squared error (RMSE) $\sqrt{N^{-1}\sum_{i=1}^{N}(\Delta_i(x))^2}$ as a function of the time lag x. The figure shows the mean of 1000 simulations, where the true time lag θ has been set to 15 units of time (minutes). Notice that the RMSE equals the standard deviation of the markup (σ) when the minimum is attained. Figure 4.7b shows the RMSE when the true time lag is zero ($\theta = 0$).

The sum of N squared standard normal random variables is known to be chi-squared distributed with N degrees of freedom (Abramowitz and Stegun, 1964, p. 940). It follows that for a fixed x, the sum $\Sigma^N \Delta_i^2$ $\Delta^2{}_i$is a scaled chi-square:

$$\sum_{i=1}^{N}\Delta_i^2 = (\sigma^2|\theta - x| + \sigma_\epsilon^2)\,\chi^2{}_N \tag{4.9}$$

4.6.2 Results of time lag analysis

Applying the time lag estimation method to our data, we find strong evidence in favor of an accurate time stamp on the average trade. For the full sample, the minimizer of equation 4.8 is zero, and the RMSE increase in accordance with the theoretical distribution laid out in equation 4.9 both backward and forward in time. Figure 4.8a show the empirical RMSE for all ∼ 257, 000 trades in the sample, for time lags and leads ranging from 1 to 30 minutes.

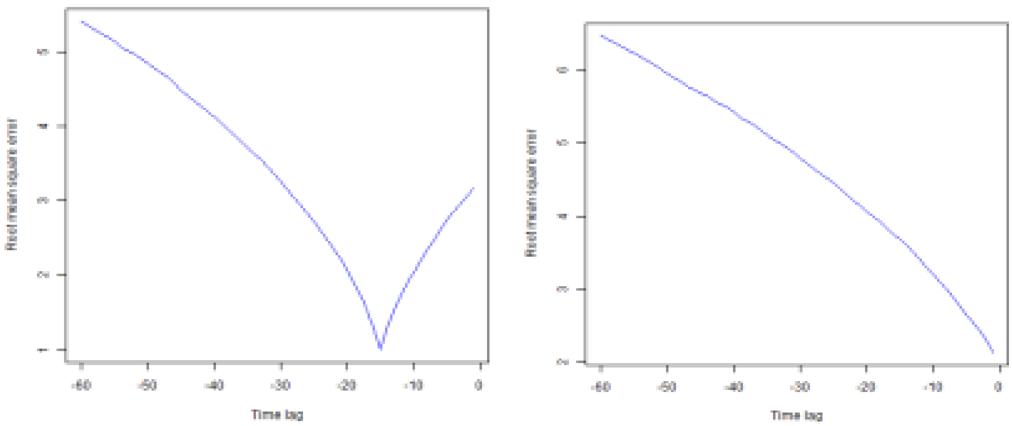
For the sample of trades involving human interaction (i.e. marked as ”direct”), the evidence is in favor of a time lag between 0 and 3 minutes, with a point estimate (minimizer of equation 4.8) of 2 minutes.

Figure 4.8b shows the empirical RMSE for these ∼ 4, 300 trades.

Figure 4.7 Theoretical RMSE Across Time Lags for Full Sample and Direct Trades

(a) Theoretical RMSE - 15 minutes lag ($\sigma = .8, \sigma_\epsilon = 1, \theta = 15$)

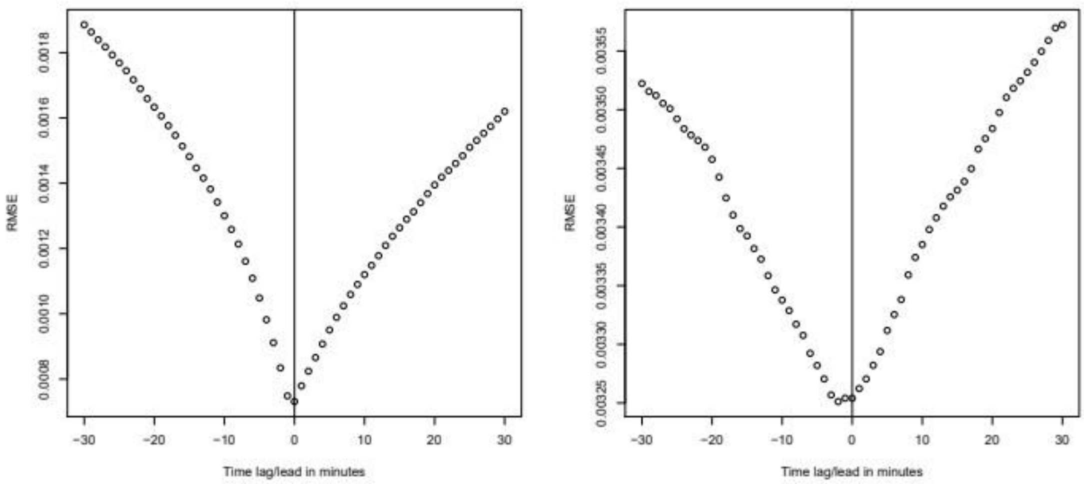
(b) Theoretical RMSE - zero time lag ($\sigma = .8, \sigma_\epsilon = 2, \theta = 0$)



In our analyses of both inventory change and price dynamics, we find evidence of pre-trade effects on a 30 minute horizon. The time lag analysis of this section show that it is highly unlikely that these effects are caused by a time lag in the bank’s booking of trades in its internal systems.

Figure 4.8 Empirical RMSE Across Time Lags for Full Sample and Direct Trades

(a) Empirical RMSE - all trades (n = 257, 241)   (b) Empirical RMSE - direct trades only (n = 4, 316)



4.7 Descriptive statistics

Table 4.4: Absolute volume by trade size.

Unit: Thousand EUR.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	N
All trades	0.00	23	120	500	500	180000	257241
Small trades	0.00	20	100	270	500	1000	239314
Medium trades	1000.00	1500	2000	3200	4000	25000	17800
Large trades	25000.00	30000	39000	58000	57000	180000	127

Table 4.5: Number of trades by venue.

	Direct	SBP	MBP 1	MBP 2	API
All trades	4316	75329	60796	8755	108045
Small trades	3223	68854	55979	6782	104476
Medium trades	1019	6442	4814	1966	3559
Large trades	74	33	3	7	10

Table 4.6: Mean trade size by venue.

Unit: Thousand EUR.

	Direct	SBP	MBP 1	MBP 2	API
All trades	2848.65	647.65	523.35	1254.76	227.84
Small trades	232.04	344.11	379.81	459.36	152.43
Medium trades	6078.35	3704.57	2170.32	3850.44	2353.51
Large trades	72338.63	37238.70	36141.54	42868.17	31592.00

Table 4.7: Customer spread by trade size.

Unit: Pips.

	1st Qu.	Median	Mean	3rd Qu.	N
All trades	-1.20	0.20	0.14	1.50	257241
Small trades	-1.20	0.20	0.14	1.50	239314
Medium trades	-1.50	0.10	0.00	1.70	17800
Large trades	-5.00	0.30	5.30	6.30	127

Table 4.8: Price change ($p_{t-30} - p_t$)) from t=-30 to 0 minutes by trade size.

Unit: Pips.

	1st Qu.	Median	Mean	3rd Qu.	N
All trades	-12.50	-2.00	-2.43	7.50	257241
Small trades	-13.00	-2.50	-2.50	7.50	239314
Medium trades	-11.00	-1.00	-0.93	10.00	17800
Large trades	-6.50	3.00	2.10	9.30	127

Table 4.9: Customer spread by venue.

Unit: Pips.

	1st Qu.	Median	Mean	3rd Qu.	N
All	-1.20	0.20	0.14	1.50	257241
Direct	-1.50	0.50	-0.20	2.50	75329
SBP	-1.00	0.20	0.26	1.50	60796
MBP 1	-1.30	0.10	0.04	1.50	108045
MBP 2	-1.30	0.30	0.14	1.70	8755
API	-1.30	0.10	0.12	1.50	4316

Table 4.10: Price change ($p_{t-30} - p_t$)) from t=-30 to 0 minutes by venue.

Unit: Pips.

	1st Qu.	Median	Mean	3rd Qu.	N
All	-12.50	-2.00	-2.43	7.50	257241
Direct	-10.00	-0.50	-0.93	8.50	75329
SBP	-13.00	-3.00	-3.40	6.50	60796
MBP 1	-12.00	-1.50	-1.30	9.50	108045
MBP 2	-10.00	-1.50	-2.20	6.50	8755
API	-13.00	-2.00	-2.40	7.50	4316

ASYMMETRIC INFORMATION AND OTC BID-ASK SPREADS

Liquidity, a critical product of any financial market, is important in itself and for its promotion of price efficiency, risk sharing, and capital formation. This paper examines how information asymmetries affect liquidity prices in OTC markets. Seminal theory from Glosten and Milgrom (1985) posits that market makers widen the spread to protect themselves from the adverse-selection risk posed by better-informed counterparties. Extensive research supports this hypothesis for order-driven equity markets (e.g., Glosten and Harris, 1988; Huang and Stoll, 1997), but in OTC markets the limited available evidence indicates that dealers quote narrower spreads to better-informed counterparties (Ramadorai, 2008; Pintér et al., 2022). This "information chasing" strategy is rational if OTC dealers can profitably exploit information gained from trades with informed clients (Naik et al., 1999; Pintér et al., 2022).

This paper highlights two additional contrasts between OTC and order-driven markets with respect to asymmetric information. Contrast 1: OTC client spreads are determined by two information asymmetries rather than one: the asymmetry between dealer and client and the asymmetry between informed and uninformed participants in the interdealer market. Contrast 2: OTC dealers employ three strategies, rather than one, when responding to client information. For some clients they narrow the spread consistent with information chasing (Naik et al., 1999; Pintér et al., 2022); for other clients they widen the spread consistent with adverse selection (Glosten and Milgrom, 1985; Easley and O'Hara, 1987); for uninformed clients they leave the spread unchanged.

We exploit the complete trading record of a major dealing institution in one of the world's most active OTC contracts, EUR-USD. Daily trading in this one contract, at roughly \$600 billion (B.I.S., 2022),[32] exceeds contemporaneous daily trading in all US and European equities (respectively \$161 bn and €41 bn) plus all US corporate and municipal bonds (respectively \$47 bn and \$13 bn).[33] Our analysis thus sheds light on a contract of global importance as well as OTC markets as a class. The data include roughly 500K trades worth close to €400 billion and are extraordinarily detailed. Most importantly, they provide an identifier and type for each counterparty, advantages that are rare in microstructure research but essential for identify-

[32] Bank for International Settlements (2022). *Triennial Survey of Foreign Exchange Turnover in April 2022*.

[33] European equities: CBOE Global Markets (2023). *European Equities Market Share by Market*. US equities: CBOE Global Markets (2023). *U.S. Equities Market Volume Summary*. US corporate bonds: Securities Industry and Financial Markets Association (2023). *US Corporate Bond Statistics*. US municipal bonds: Securities Industry and Financial Markets Association (2023). *US Municipal Bond Statistics*.

ing the influence of OTC information asymmetries.

Contrast 1, the relevance of two information asymmetries for OTC client spreads rather than one, arises from the two-tier structure of OTC markets. In the interdealer tier, dealers trade with each other and, in FX, with sophisticated non-dealers under prime brokerage contracts – an expansion that motivates the broader label, "core market." In the second tier, dealers provide liquidity to clients, setting each client quote as a markup relative to the prevailing same-side core-market price. By construction, therefore, a client half-spread is the sum of two components, each of which is influenced by a different information asymmetry. The core-market half-spread is influenced by asymmetric information among dealers; the client-specific markup is influenced by asymmetric information between dealer and client.

We first estimate the familiar Huang and Stoll (1997) spread decomposition model using prices from the bank's liquidity-providing trades on core-market order-driven platforms, quotes from which are visible to the entire professional trading community. The regression distinguishes traditional interdealer matching engines, where all quotes are firm, from newer platforms offering "last-look" optionality. With last look, liquidity providers can decline to trade during a fraction of a second (typically 80 -100 milliseconds) after a counterparty accepts a quote (Phillips et al., 2020; Chaboud et al., 2023). Last-look platforms are most relevant to our analysis because their spreads are relatively narrow so their quotes will generally serve as the 'prevailing market price' on which client prices will be based. Results indicate that the last-look core-market spread is 25% of the mean client spread, a share that rises to 35% to 74% for the clients that are most sensitive to execution costs: hedge funds, commodity trading firms, brokers, and client banks.

Three OTC-specific market features also explain Contrast 2, the multiplicity of OTC dealer responses to client information. First, the structure of the OTC trading game precludes client anonymity: OTC clients move first by requesting quotes from a specific dealer. OTC dealers can therefore condition client spreads on client characteristics including information (Ramadorai, 2008). Second, dealer-client relationships are typically long-lasting (Hendershott et al., 2020; Hau et al., 2021), so dealers have ample time to assess the client's information. Third, OTC dealers have instantaneous market power due to clients' non-trivial search costs (Duffie et al, 2005).

Client information inevitably creates adverse-selection risk for dealers, which could motivate them to widen the spread for better-informed clients (Glosten and Milgrom, 1985; Easley and O'Hara, 1987). In the OTC context, however, client information can also bring benefits: dealers can use it to take parallel proprietary positions in the core market (Naik et al., 1999) or to revise their quotes to other agents (Pintér et al., 2022). If the benefits of client information dominate, dealers will attract the repeat business of better-informed clients by quoting them narrower spreads (Naik et al., 1999; Pintér et al., 2022). If benefits and costs are roughly equal, or if the client is uninformed, client information will have no influence on client spreads (Naik et al., 1999).

To evaluate dealer responses to client information we develop a comprehensive markup decomposition model tailored to the OTC setting. The model allows dealers to price discriminate in response to client information and three additional client characteristics: trading activity (Bernhardt et al., 2005; O'Hara et al., 2018); sophistication (Duffie et al., 2005; Green et al., 2007); and access to additional counterparties (Hendershott et al., 2020; Hau et al., 2021). The model allows dealers to vary their strategic response to client information across client types. Client information is measured as mean signed post-trade returns following Anand and Subrahmanyam (2008). As is standard, the model also includes order-handling and inventory costs (Glosten and Harris, 1988; Huang and Stoll, 1997). Finally, the model incorporates the possibility that dealers smooth spreads over time within each client relationship. This enhances a dealer's reputation for reliability and minimizes the irritation and confusion that customers experience when prices change frequently (Dholakia, 2016). The model is estimated using GMM with robust standard errors clustered by date. The model's explanatory power is strong at 0.58 and estimated coefficients consistently support the model's motivating theories. Findings are sustained over multiple robustness tests.

Results indicate that dealers adopt all three of the potentially rational responses to client information. They narrow the spread for better-informed brokers, consistent with information chasing. They widen the spread for better-informed hedge funds, commodity trading firms, and client banks, consistent with adverse selection. They make no information adjustments for non-financial clients, private clients, and low-leverage asset managers ("real-money funds") such as pension funds or mutual funds, which appear to be uninformed.

The paper finishes by investigating the choice between information chasing and adverse selection. We hypothesize that the choice is driven in part by HFT, which reached FX in the early 2000s. Until then FX dealing institutions had consistently sought to attract the business of better-informed clients (Clyde, 1996; Ramadorai, 2008), but they soon began expressing concern (Markets Committee, 2011). HFT strategies such as latency arbitrage essentially add a new dimension of adverse selection, especially for dealers whose quote updating systems could not keep up with HFT (Menkveld and Zoican, 2017). We hypothesize that trades with HFT clients bring relatively high costs and low benefits, motivating wider spreads and a greater sensitivity of spreads to client information. We test this hypothesis by re-estimating the markup decomposition model with hedge funds partitioned into those that do and do not engage in HFT. Consistent with the HFT hypothesis, dealers widen spreads with client information three times more aggressively for HFT hedge funds than for other hedge funds.

This paper contributes to the vast literature on liquidity pricing in financial markets. [34] Research on OTC spreads was relatively infrequent until around 2000 and research on information asymmetries in OTC markets remains rare. We are the first, to our knowledge, to discuss and evaluate the relevance of core-market information asymmetries for OTC client spreads. Early evidence relevant to dealer-client information asymmetries led researchers to question the relevance of adverse selection. OTC client spreads are inversely related to trade size (e,g., Reiss and Werner, 1996; Bernhardt et al., 2005; Edwards et al., 2007), though informed agents could rationally choose to make large trades (Easley and O'Hara, 1987); FX spreads are narrower for financial than non-financial clients, though the former are better informed (Osler et al., 2011; Bjønnes et al., 2021); finally, the predictions of adverse selection are not fulfilled when standard spread decomposition models are estimated with OTC data (Osler et al., 2011).

The implications of these findings for adverse selection are unclear, however. The inverse relation between spreads and trade size, for example, could be explained by fixed order-processing costs or the dealers' efforts to attract future business from active clients (Bernhardt et al., 2005). Disentangling the many influences on OTC spreads requires data with client identifiers, which have been elusive. Ramadorai (2008), the first study to benefit from such data, provides evidence that FX dealers narrow spreads for better-informed asset managers. Pintér et al. (2022) provide further evidence for that strategy in the UK government bond market. The wide range of clients in our data enable us to extend this line of inquiry by testing whether dealers vary their response to client information across client types. Our findings resolve the earlier puzzles by showing that OTC dealers employ all three of the available strategies for responding to client information.

The rest of this paper has five sections and a conclusion. Section 5.1 describes our data. Section 5.2 analyzes the contributions of asymmetric information in the FX core market to client execution costs. Section 5.3 develops the markup decomposition model we use to identify the dealers' strategic responses to client information. Section 5.4 presents the results of estimating

[34] E.g., Edwards et al. (2007); Goldstein et al. (2007); Green et al. (2007); Harris and Piwowar (2007); Hotchkiss et al. (2008); O'Hara et al. (2018); Hendershott et al. (2020); O'Hara and Zhou (2021); Hau et al. (2021); Feldhütter and Poulsen (2018); Friewald and Nagler (2024); Collin-Dufresne and Fos (2015); Collin-Dufresne, et al. (2020).

that model and a set of robustness tests. Section 5.5 examines dealers' choice between information chasing and strategic dealing. Section 5.6 concludes.

5.1 Data

Our data comprise the complete front-office record of spot trades in EUR-USD at a top-10 dealing bank over the 68 trading days from 2 January to 20 April 2012. Trading in FX, arguably the world's largest OTC market, averages $7.5 trillion daily (Bank for International Settlements, 2022). Trading in EUR-USD takes place non-stop from around 12 am GMT on Mondays to around 21 GMT on Fridays and at lower volumes on weekends. Our trade data include the standard fields – transaction price, quantity, date, and time (to the second) – plus trade initiator, trade direction, trading platform, counterparty ID, many client types assigned by the bank, exact markup for client trades, a flag for trades priced by human interbank dealers, and both prime broker and prime broker client when relevant. All trades have complete information. Client disaggregation exceeds that in previous FX research (e.g., Osler et al., 2011; Menkhoff et al., 2016; Ranaldo and Somogyi, 2021) and proves valuable. Our evidence, for example, reveals major differences among the hedge funds, real-money funds, and private clients, all of which are typically grouped together as "financial clients" or "institutional investors."

Our sample period includes a fairly important event, the resolution of private-sector contributions to a major Greek debt bailout. After lengthy negotiations the Greek terms were announced on 21 February; the European Commission entered a memorandum of understanding with the Greek government and the Bank of Greece on 1 March; the bond exchange took place in April (Bank for International Settlements, 2020). Fortunately, these events did not generate unusual exchange-rate dynamics that could compromise our analysis. Figure 5.1 presents exchange rates and realized volatility (20-day, centered) during our sample period together with the preceding and subsequent 12 months. EUR strengthened early in the sample period, as the memorandum of understanding approached, and then stabilized; its sample mean of $1.31/€ is close to its longer-period mean, $1.33/€. Realized volatility was moderate throughout the sample period and its sample mean, 2.66%, is likewise close to the longer-period mean, 2.69%.

5.1.1 Core-market Trades

Core-market trades are exchanges among members of the self-described "professional trading community," a group we specify as the top 50 *Euromoney* dealing banks plus agents trading via prime broker agreements. The FX core market was, until the early 1990s, a quote-driven interdealer market similar to those today in less liquid assets such as US municipal bonds (Harris and Piwowar, 2006), precious metals, CDSs (Collin-Dufresne et al., 2020), or other OTC derivatives (Hau et al., 2021).[35] Order-driven trading quickly came to dominate after EBS and Reuters introduced matching engines in the early 1990s. In the late 1990s a few large and sophisticated hedge funds were allowed to participate under prime brokerage contracts, a privilege subsequently extended to proprietary trading firms as well. Since around 2000 the available core-market trading platforms have expanded to include newer matching engines and dark pools, and quote-driven trading remains an option. During our sample period EBS and Reuters Matching enforced strict price-time priority, prohibited hidden orders, and set minimum tick sizes at 0.5 pip (EBS) and 1 pip (Reuters). (A pip or "price improvement point" is $0.0001 in EUR-USD.) The newer matching engines generally offer last-look protection, hidden orders, and decimal-pip pricing. Our bank acts as a liquidity provider for the vast majority (89%) of its trades with core-market participants.

We base our empirical analysis of the core market on order-driven trades because their transparency and broad dissemination enable them to serve as the foundation for client quotes. As explained by the Municipal Securities Rulemaking Board (MSRB, 2018), OTC client prices are set relative to the "prevailing market price." Core-market trades executed in dark pools or quote-driven settings are opaque to all but the immediate counterparties, which prevents them from serving this function. Our sample includes 131,269 core-market order-driven trades, worth in total €117 billion (Table 5.1). These trades are roughly evenly split between traditional interdealer platforms and last-look platforms but individual trades over EBS and Reuters tend to be larger so they account for three-quarters of total value exchanged.

5.1.2 Client Trades

Our sample includes 2,388 clients that made 257,241 trades worth €126.0 billion in aggregate.[36] The sample's 1,229 financial clients include 43 hedge funds, 19 commodity trading firms, 781client banks, 91 brokers, 257 real-money funds, and 88 private clients. Commodity trading firms (CTFs) engage in proprietary trading, like hedge funds, but focus on commodity futures contracts (Pirrong, 2015). Client banks are small or regional and generally focus on client service; a few niche banks use HFT. Some of the brokers offer narrow spreads to retail traders via online platforms and trade with our bank to manage inventory. Other brokers provide advice and execution services for institutional clients. Real-money funds are low-leverage asset

[35] Unlike many of those OTC markets, however, dealers in liquid currency pairs never act as brokers.
[36] We exclude the handful trades in which a client provided liquidity to our bank over an order-driven platform.

managers such as pension funds, mutual funds, and endowments. Private clients manage the assets of wealthy individuals or families. Among non-financial clients, 169 are MNCs, defined as a firm with a specialized treasury unit or at least €100 million in annual sales; the SMEs, of which there are 990, have annual sales below €100 million and no treasury unit.

Table 5.2 provides descriptive information for clients. The mean markup for client trades is 0.44 pips or 0.34 bps at the sample's mean exchange rate of $1.31/€. Mean markups vary widely across client types; they are essentially zero for brokers and reach 38.4 pips for private clients. Trade frequency also varies widely. The average real-money fund, private client, MNC, or SME trades at most once or twice per month; the average hedge fund or client bank trades once or twice daily; the average CTF or broker trades over 15 times daily.

The bank we study, like most large dealing banks, internalizes the vast majority of its client trades and prices them electronically. Human dealers set the quotes for the remaining trades, either in response to a client's request or because the bank flags the trade as potentially toxic. Whether priced by an algorithm or by a human, the client price is set as a markup relative to a composite same-side core-market price constructed by the bank's pricing engine, which uses input from several order-driven platforms. The markup is tailored to the individual client, the behavior and preferences of which are digitally scrutinized by a team of specialists and familiar to their human salespeople.

5.1.3 Client Trading Platforms

Clients of our bank trade on more than 20 platforms that we partition into four groups: direct trading, single-bank platforms (SBPs), multi-bank request-for-quote platforms (MRFQs), and application programming interfaces (APIs).

For a <u>direct trade,</u> the client reaches out to a salesperson by telephone, email, txt, etc., to request quotes. Just 2% of trades are handled directly, presumably because this approach uniquely requires the costly time of a human dealer (Bjønnes et al., 2025). Many clients adopt this approach, nonetheless; all trades by private clients and over two-thirds of SME trades are handled directly.

An <u>SBP</u> is a desktop interface over which clients can request quotes from, and trade with, a specific bank. The simplest SBPs allow the client to place orders in advance for trades at a daily fix; the most complex provide click-and-trade functionality with streaming actionable quotes. Client banks and real-money funds rely most heavily on SBPs.

On an <u>MRFQ</u> platform, the client solicits simultaneous quotes from multiple dealing institutions (Hendershott and Madhavan, 2015). MRFQs enhance competition among dealers, which can bring narrower spreads, but they also reveal information about a client's intended trades. The installation and maintenance of MRFQ platforms requires considerable technical sophistication and usually takes at least two months. MRFQs are used most frequently by hedge funds, CTFs, real-money funds, and MNCs.

An API allows client algorithms to trade electronically at prices streamed by the dealer and requires by far the most technological sophistication from clients. These platforms are used for HFT by hedge funds and CTFs and for efficient inventory management by the most active brokers.

5.2 Client Spreads and Core-Market Information Asymmetries

This paper's first contribution is to highlight and estimate the influence on OTC client spreads of information asymmetries among professional traders. An OTC client half-spread is the sum of the client-specific markup plus the core-market half-spread and thus includes the adverse-selection component of the core-market spread. Core-market adverse selection risk could be substantial given the wide variation in information among dealing institutions. Bjønnes et al. (2021) provide evidence that an FX dealing bank's information rises monotonically with its size and with the extent of its trading with hedge-fund clients. They also provide evidence that much of a bank's information is self-generated, consistent with Evans and Lyons' (2012) hypothesis that dealing banks infer currency fundamentals from patterns in client trades.

5.2.1 Regression Model

To estimate the adverse-selection component of core-market spreads we apply the familiar Huang and Stoll (1997) spread decomposition model, the assumptions of which we summarize briefly. The price for trade t, P_t, is determined as market mid-quote, M_t, plus or minus an effective half-spread, $S/2$, plus an i.i.d. mean-zero noise term, h_t: $P_t = M_t + \frac{S}{2}D_t + \eta_t$. Trade direction, D_t, is +1 (-1) when the bank's counterparty buys (sells) the base currency, EUR. M_t is the asset's perceived true value, V_t, adjusted for price shading in response to inventory imbalances: $M_t = V_t - \beta\frac{S}{2}(I_t - I^*)$, where I_t is current inventory and I^* is desired inventory. The asset's perceived true value, V_t, is updated in response to the information in the most recent trade, $\alpha\frac{S}{2}D_{t-1}$, and an i.i.d mean-zero public news shock, ε_t: $V_t = V_{t-1} + \alpha\frac{S}{2}D_{t-1} + \varepsilon_t$. Together, these assumptions imply the following expression for the current traded price:

$$P_t = V_{t-1} + \alpha\frac{S}{2}D_{t-1} - \beta\frac{S}{2}(I_t - I^*) + \varepsilon_t + \frac{S}{2}D_t + \eta_t. \quad (5.1)$$

V_t, the asset's true value, is unknown so the model uses the return, $\Delta P_t = P_t - P_{t-1}$, as dependent variable:

$$\Delta P_t = \frac{S}{2}(D_t - D_{t-1}) + \alpha\frac{S}{2}D_{t-1} - \beta\frac{S}{2}\Delta I_t + e_t, e_t \equiv \eta_t + \varepsilon_t - \varepsilon_{t-1}. \quad (5.2)$$

We estimate the model on the order-driven trades in which our bank provided liquidity to other core-market participants: core-market RFQ trades are excluded because their prices are not fed into our bank's pricing engine. As recommended by Huang and Stoll (1997), we identify a single trade as the sum of all transactions at a given moment with the same time stamp and price. By including trade direction instead of trade size, the model implicitly assumes that all trades have roughly the same size. To achieve greater consistency in trade sizes our main regressions exclude trades below €250,000. The coefficients S, a, and b are estimated separately for firm-quote (FQ) and last-look (LL) platforms.

5.2.2 Main Results

As shown in in Table 5.3, column 1, the estimated effective core half-spread is 0.14 pip for last-look platforms and (0.61 pip) for firm-quote platforms.[37] The difference between last-look and firm-quote platforms is highly significant. The compensation for adverse-selection risk is 0.12 pip on last-look platforms and 0.37 pip on firm-quote platforms, a difference that accounts for over half of the difference in effective half-spreads. The difference in adverse-selection components is consistent with the absence of the last-look option on EBS and Reuters Matching and provides reassurance that the model is indeed capturing adverse selection, which need not always be the case (Van Ness et al., 2001; Collin-Dufresne and Fos, 2015). The proportionate contribution of adverse-selection to effective core spreads is 90% for last-look platforms and 61% for firm-quote platforms. These shares are relatively large compared to estimates for large US stocks generated with the same methodology, where the average is 46% and the range across stocks is 21% to 64% (Huang and Stoll, 1997). The importance of information asymmetries for core FX spreads is unsurprising because, as shown by Bjønnes et al. (2021), the most active FX dealers have a highly-profitable information advantage relative to the least active dealers.

The adverse selection component of spreads on last-look platforms is 22% of the mean client half-spread in our data, 0.58 pip,[38] though this proportionate contribution varies widely. It is 5% or less for clients that trade infrequently – real-money funds, private clients, MNCs, and

[37] The interpretation of coefficient magnitudes as pips is appropriate given the median core-market trade size of €1 million on both platform types.

[38] The mean client half-spread is the mean client markup, 0.44 pip, plus the 0.14-pip effective last-look core half-spread.

SMEs – but it ranges from 31% to 84% for clients that trade frequently, specifically hedge funds, CTFs, client banks, and brokers.

The adverse-selection component of spreads on firm-quote platforms, 0.37 pip, conforms to the conception of adverse-selection risk embedded in Glosten and Milgrom (1985) and subsequent models. The difference between this and the adverse-selection component on last-look platforms, 0.14 pip, represents the adverse selection costs avoided by the last-look option. If last look were unavailable and core-market half-spreads included this additional cost, core-market adverse selection would represent 35% of mean half-spreads across all clients and 43% to 61% of mean half-spreads for clients that trade frequently. These as-if costs for firm-quote platforms are potentially most relevant to other OTC markets, where last-look optionality is generally unavailable.

The conclusion that core-market adverse selection has a substantial influence on OTC client spreads is robust. The estimates of S and a change little when the sample is limited to trades during relatively the active period of 8 GMT to 22 GMT (Ito and Hashimoto, 2006) (Table 5.3, column 2). When the sample includes trades below €250,000 the effective half-spread roughly doubles for last-look platforms but changes little for firm-quote platforms (Table 5.3, column 3). This reflects the concentration of tiny trades on last-look platforms: the average trade size falls substantially for last-look platforms but not for firm-quote platforms. These adjustments do not affect our qualitative conclusions regarding the importance of core-market adverse selection for client spreads. Core-market adverse selection on last-look platforms continues to account for a substantial share of mean client half-spreads, 25%, and that share ranges from 35% to 74% for hedge funds, CTFs, client banks, and brokers. Our qualitative conclusions are also sustained when price shading is excluded with the original sample (Table 5.3, column 4).

It is notable that the price shading coefficients are small but statistically significant, which suggests a shift over time in FX dealer behavior. Despite the abundance of evidence for price shading in equity markets (e.g., Huang and Stoll, 1997), studies of FX markets have found, instead, that dealers offload inventory positions quickly via outgoing interdealer trades (Bjønnes and Rime, 2005). Dealers explain this practice in terms of transparency risk: pricing shading conveys information about the dealer's position that could be exploited by other traders. This apparent shift towards price shading on order-driven platforms could be related to the roughly-contemporaneous introduction of platforms that permit hidden orders and reduce information leakage.

5.3 Asymmetric Information Between Dealers and Clients: An Empirical Model

This paper's second contribution is to provide evidence that OTC dealers adopt three different strategies in response to client information. Consistent with information chasing, dealers narrow the quotes for better-informed brokers. Consistent with adverse selection, dealer widen quotes for better-informed hedge funds, CTFs, and client banks. Dealers make no quote adjustments related to client information for real-money funds, private clients, MNCs, and SMEs.

This section presents our empirical model of dealer responses to client information. The next section presents the results of estimating that model on our data. Each observation is a client trade, indexed by *t*. The dependent variable is the dealer's chosen markup, *Markup*, the precise value for which is provided by the bank. The model controls for all OTC spread determinants highlighted in the literature as relevant to a single FX contract plus intertemporal smoothing, a dealer strategy introduced here.

$$Markup = \sum_j InfoCT'_j\pi_j + \beta TV + Pltshr\delta_2' + \theta Skew + PCP\psi'$$

$$\varphi_0 + Time_t\,\kappa_1' + (\lambda_1 + \varphi_1)Sz + DT\,(\varphi_2 + \varphi_3 Sz) + (\kappa_2 + \lambda_2)HL + (\kappa_3 + \lambda_3)\,CSprd + \zeta\,, \qquad (5.3)$$

$$j = \{HF, CTF, ClientBank, Broker, RM, PC, MNC, SME\}.$$

The first line of Equation (5.3) includes variables to capture client information and other relevant client characteristics: trading activity (*TV*), sophistication (*Pltshr* and *Skew*), and the client's access to other counterparties (*PCP*). The second line includes variables intended to capture order-handling and inventory costs plus smoothing: indicators for segments of the trading day, *Time*; trade size *Sz*; volatility, *HL*; and the contemporaneous core-market spread, *CSprd*.

5.3.1 Client Information

The first term in Equation (5.3), $\sum_j InfoCT'_j\pi_j$, captures the dealer's response to client information, *Info*. The coefficient, $\pi_{j(c)}$, is allowed to vary by client type and will be negative under information chasing (Naik et al., 1999; Pintér et al., 2022); positive under adverse selection (Glosten and Milgrom, 1985; Easley and O'Hara, 1987); and otherwise insignificant. $Info_c$ is measured as client *c*'s mean signed post-trade return, following Anand and Subrahmanyam (2008): $Info_c \equiv \text{Mean}[D_{tc}(m_{t+k} - m_t)]$. D_{tc} is the direction of trade *t* if the counterparty is client *c*; m_t is the near-contemporaneous mid-quote taken from a complete sample of best bid and ask quotes taken from the Reuters Dealing 3000 platform, which provides mid-quotes at evenly-spaced time intervals (unlike our irregularly-spaced transaction data); *k* is either 1 minute or 30 minutes. Post-trade returns in FX can be measured accurately due to the high trading volume. In less active markets, such as those for most U.S. corporate and municipal bonds, trades in a given bond occur infrequently, so it is challenging to estimate the asset's market price after a fixed post-trade interval.

Table 5.4 reports descriptive statistics for $Info_c$ by client type. Mean 1-minute post-trade returns are positive and significant for hedge funds and brokers, consistent with the hypothesis that their trades are informed on average, but not for other client types. Post-trade returns for hedge funds and brokers would generate annual gains of 2.0% and 3.6%, respectively, at the average trade frequencies of those client types. We partition trades into four size categories to investigate whether informed trades are large, as predicted in Easley and O'Hara (1987), or medium-sized, as predicted by the stealth trading hypothesis (Barclay and Werner, 1993). The categories are: tiny ∈[€0,€0.5); small ∈[€0.5, €1.0); medium ∈ [€1.0,€ .05); large ∈ [€5.0,€∞] (all amounts in million EUR). For 1-minute post-trade returns there is no evident relation between trade size and information. These statistics suggest, however, that trades by CTFs and client banks in the top three size categories carry information, despite the insignificance of their overall average post-trade returns, a finding that proves relevant in Section 5.4. For 30-minute post-trade returns standard errors are far larger and the overall value is only significant for hedge funds. The disaggregation by trade size is not presented because only two of the 32 entries are significant, specifically those for tiny and small hedge-fund trades.

5.3.2 Other Client Characteristics

The model includes three additional client characteristics: trading activity, sophistication, and access to additional counterparties.

Trading activity, TV, is measured as the client's log trading volume, TVc ≡ Ln(TrdVolc). Bernhardt et al. (2005) shows that rational dealers will quote narrower spreads to active clients, to attract their repeat business; we therefore expect □<0. Evidence for OTC activity discounts is provided by Bernhardt et al. (2005) for the LSE, Hau et al. (2021) for FX derivatives, and O'Hara et al. (2018) for U.S. corporate bonds.

The variables *PltShr* and *Skew* should capture "sophistication," meaning the negotiating skills, market knowledge, and access to technology required to minimize execution costs (Green et al., 2007). Sophisticated clients are more sensitive to spreads, which limits the dealer's market power and leads to narrower spreads in equilibrium (Ho and Stoll, 1981; Duffie et al., 2005;

Green et al., 2007). The most sophisticated clients tend to be hedge funds, which often hire professional FX traders, and client banks. The least sophisticated tend to be private clients and SMEs, some of which are not aware that FX dealers trade as principals under the obligation to maximize shareholder welfare (Galanek, 2010) rather than as agents under the obligation to maximize client welfare, like stock brokers.

For trade *t* by client *c*, *PltShr* includes the share of *c*'s trades handled directly, over SBPs, and over MRFQs (the API share is omitted). In using platform shares to capture sophistication we follow Liu et al. (2023), a choice that can be clarified by considering a client that needs to exchange currencies for the first time but has no in-house expertise. It will logically reach out to its relationship bank, which will identify whom to contact for a direct trade. Direct trades require the time of a human dealer and therefore incur higher order-handling costs than all-electronic alternatives. Nonetheless, clients that trade infrequently may continue to rely on direct trading indefinitely because they have little incentive to learn whether execution costs can be reduced and how to do so: that is, remaining unsophisticated can be a rational choice. Consistent with this portrait, 100% of private clients, all of which trade once per week or less, rely on direct trading, as do 96% of SMEs that trade once per week or less. Further, almost none of these private clients or SMEs rely on SBPs (0% and 3%, respectively) or MRFQs (0% and 2%, respectively).

Larger clients that trade infrequently, and especially those with a Treasury function, are more likely to have in-house expertise and thus less likely to rely on direct trading: just 62% of MNCs that trade once per week or less rely on direct trading. Unlike SMEs, many MNCs that trade infrequently rely on SBPs and MRFQs (26% and 24%, respectively) (Figures 5.2A-5.2C). These differences between small and large clients are more pronounced among financial clients, which are more likely to have in-house expertise. Among financial clients other than private clients ("other financial clients") that trade once per week or less, just 22% trade directly while 40% rely on SBPs and 44% rely on MRFQs.

Clients that trade frequently have strong incentives to learn about the alternatives to direct trading and most do so. Some learn about SBPs from their relationship bank as it attempts to retain their business; others learn through their own investigation, and most clients that trade frequently wisely avoid direct trading. Consistent with this portrait, Figure 5.2A shows falling shares of clients that rely on direct trading generally falls as trade frequency increases beyond once per week. In parallel, SMEs and other financial clients shift towards SBPs (Figure 5.2B) while MNCs shift towards MRFQs (Figure 5.2C). For active clients of all types, half or less rely on direct trading and a substantial majority rely on either SBPs or MRFQs.[39]

Skew is intended to capture the dealers' tendency to skew quotes for unsophisticated clients with predictable trade direction. For clients that are likely to buy (sell), the dealer quotes a normal spread while raising (lowering) both quotes relative to the mid-quote. A trade in the client's normal direction generates an above-average markup but unsophisticated traders, like the busy SME employee who spends 99% of their time on other responsibilities, may be unaware. $Skew_c \equiv ShrDT_c BuySell_c$. is the interaction of two terms: $ShrDT_c$, the client's share of direct trades as an inverse measure of sophistication, and $BuySell_c \equiv |\#Buys_c – Sells_c|/(\#Buys_c + \#Sales_c)$, the predictability of the client's trade direction ($BuySell_c \in [0,1]$).

Access to Potential Counterparties. The variable *PCP* is included to capture cross-client variation in access to potential counterparties. It comprises indicator variables for clients with unusually many or unusually few potential counterparties. Having multiple counterparties enables a client to decline to trade, and seek a better price with another dealer, when quoted an unattractive spread. One can interpret access to other counterparties as a high price elasticity of demand (Ho and Stoll, 1981), low search costs (Duffie et al., 2005), or strong negotiating leverage (Green et al., 2007). Regardless, it should bring narrower spreads in equilibrium, evidence for which is provided in Hendershott et al. (2020) and Hau et al. (2021).

Though our data do not explicitly list the number of a client's FX dealing relationships, they nonetheless provide relevant information. They indicate, for example, clients that trade via prime broker, a status that provides the largest possible number of potential FX counterparties – the entire professional trading community. The first indicator in *PCP* identifies such clients and we expect $\psi_{PB} < 0$.

The data also help identify clients with few potential counterparties, a choice that is reportedly common and can be fully rational. Having just one FX dealing relationship maximizes activity discounts, as noted by O'Hara and Zhou (2021). Further, it allows the client to avoid the burden of qualifying for a line of credit at an additional dealing bank, which requires posting a large, permanent deposit or demonstrating credit-worthiness. At small firms, where processes and procedures are not well-established and employees wear multiple hats, it can be challenging to gather and present the required information. The reluctance of small firms to incur such costs

[39] According to market participants, the MNCs' preference for MRFQs often reflects internal control mandates that require bids from multiple service providers. MRFQs can generate narrow spreads by intensifying competition among liquidity providers, as noted in Section 5.2. However, a representative of a major MRFQ reports that these platforms are generally not cost-effective for clients that trade less than €100 million annually due to their high set-up costs. The process generally requires at least two months and specialized expertise.

is suggested by the striking difference between SMEs and MNCs in the platforms to which they switch as trade frequency increases. Most MNCs switch to MRFQs, which clients use to exploit multiple dealing relationships; SMEs, however, switch to SBPs instead (Figures 5.2B and 5.2C). The second indicator in *PCP* identifies SMEs that do not trade via MRFQ platforms; the third indicates private clients, none of which trade via MRFQ platforms; we expect ψ_{Priv}, $\psi_{SME} > 0$.

The major dealing banks are generally reluctant to trade banks from countries with low financial development (LFD). A dealing bank's concerns about foreign financial crises and settlement risks can be overcome, however, when it serves as correspondent bank for the foreign bank. In this case the foreign bank must maintain non-trivial deposits with the dealing bank, deposits that largely eliminate settlement risk and generate profits for the dealing bank. Figure 5.2D provides evidence that LFD client banks are indeed challenged to establish relationships with other FX dealing banks. We define an LFD country as one with 2012 IMF Financial Development Index in the lowest one-third of the distribution. The overall share of LFD banks that rely on MRFQs, 13.8%, is statistically below the overall share for other banks, 50.4%; indeed, MRFQ reliance is lower for LFD banks than other banks at every trade frequency. (Results are almost identical if the LFD cutoff is at 0.25.) The fourth indicator in *PCP* identifies LFD client banks; we expect $\psi_{LFD} > 0$.

5.3.3 Remaining Control Variables

The model's remaining variables are intended to capture order-handling, inventory costs, and intertemporal smoothing.

Most of the costs of running a trading room – employee compensation, space, technology, information resources, etc. – are fixed. The fixed cost for all-electronic trades should be captured by $\varphi_0+\varphi_1 Sz$. Trade size, $Sz_t = ln(TrdSz_t)$, is included because larger trades allow market makers to cover a fixed cost with a smaller markup; we expect $\varphi_0>0$ and $\varphi_1<0$. The additional fixed cost for direct trades is captured by $(\varphi_2+\varphi_3 Sz)DT$, where *DT* is an indicator for direct trades;

Inventory costs can theoretically include both carry and risk, but carry is negligible in spot FX because dealers do not generally hold positions overnight (Bjønnes and Rime, 2005). The contribution of inventory risk to the markup is $\lambda_1 Sz + \lambda_2 HL + \lambda_3 CSprd$. Inventory risk rises with trade size, *Sz*, and volatility, *HL* (e.g., Ho and Stoll, 1981; Bessembinder, 1994; Bollerslev and Melvin, 1994). *HL* is the high-low range of the previous trading hour; results are robust to using realized volatility. Inventory costs also depend on illiquidity, which we capture with the contemporaneous core spread on the Reuters Dealing platform, *CSprd*. We expect λ_1, λ_2, and $\lambda_3>0$.

We hypothesize that dealers smooth spreads over time within each client relationship. Dealers know that an OTC client's search for best execution is not random, undirected, or non-strategic. Instead, clients choose among dealers on the basis of a list of qualities that includes spreads plus promptness, friendliness, perks such as gala gatherings, and reliability. Smoothing enhances an impression of dealer strength and reliability. In the dealer survey of Cheung and Chinn (2001), dealers explain that "the ability to consistently offer quotes with … conventional spreads in a hectic market is regarded as an essential characteristic of a market leader" (p. 444). Two-thirds of responding dealers usually choose the spread's conventional level when quoting to each other; the profitability of a given trade is secondary. Smoothing has additional benefits: it reduces price noise that can mask a dealer's competitiveness and it spares clients the confusion and annoyance of frequent price changes (Dholakia, 2016, p. 1).

The model captures the dealer's effort to smooth spreads as follows:

$$Time\kappa_1'+\kappa_2 HL+\kappa_3 CSprd_{..}$$

Time comprises three time-of-day indicators intended to capture any tendency of dealers to move markups inversely to the intraday seasonals in core spreads (Ito and Hashimoto, 2006). Core spreads are lowest during European hours, 8–16 GMT, when volume is highest, and highest during the overnight period, when volume is lowest, 22–02 GMT; the other two time periods correspond to New York trading, 17 GMT–21 GMT (Asian hours are excluded) (Ito and Hashimoto, 2006). Smoothing implies, $\kappa_{1-Europe}>0$ and $\kappa_{1-Overnight}<0$. Smoothing implies $\kappa_2<0$ because the core spread rises with volatility (Bessembinder, 1994; Bollerslev and Melvin, 1994). Smoothing also implies $\kappa_3<0$.

The model is estimated using GMM with Newey-West standard errors clustered by date. We pause to consider whether client characteristics are highly correlated, which would create the risk of multicollinearity. Privately-informed clients have strong incentives to trade actively; active clients have strong incentives to invest in sophistication. We calculate the pairwise correlations across clients for private information ($Info_c$), trading activity (TV_c), and each client's share of direct trades ($ShrDT_c$) as an inverse scalar measure of sophistication. All prove to be moderate: $\rho(Info_c,ShrDT_c) = 0.006$; $\rho(Info_c,TV_c) = -0.0007$; $\rho(ShrDT_c,TV_c) = -0.40$.

5.4 Asymmetric Information Between Dealers and Clients: Results

The model's estimated coefficients indicate that dealers adopt all three of the potentially rational strategies for dealing with client information. The *Info* coefficient is negative and significant for brokers, confirming existing evidence for the importance of information chasing (Ramadorai, 2008; Pintér et al., 2022): the 91 brokers in our sample account for 58% of trades and 30% of trade value. A one-standard-deviation decrease in a broker's *Info* (0.74 pip) is estimated to raise its markup by 0.51 pips.

The regressions also indicate that adverse selection influences OTC spreads. *Info* coefficients are positive and significant for hedge funds, CTFs, and client banks, which jointly account for 33% of clients and 67% of client trade value. A one-standard-deviation increase in *Info* is estimated to raise the hedge fund markup by 0.49 pip, the CTF markup by 0.96 pip, and the client-bank markup by 0.31 pip. The *Info* coefficient may be relatively small for client banks because they are relatively unlikely to be informed; most such banks focus on client service and trade with our bank primarily to manage inventory. In addition, those banks' own clients tend to be uninformed because the best-informed clients prefer to trade with the dominant banks (Bjønnes et al., 2021).

This is the first rigorous evidence for adverse selection in OTC markets of which we are aware. We attribute the emergence of this evidence for adverse selection, in part, to the comprehensive set of clients represented in our data and to the comprehensive inclusion of spread determinants in our estimating model. As shown in Bjønnes et al. (2025), the originally-puzzling inverse relation between OTC spreads and client information is largely due to differences in client sophistication and potential counterparties, characteristics that cannot be identified without client identifiers. Sophisticated clients such as hedge funds are indeed better informed than average, yet they pay narrower markups because they are relatively good negotiators and choose to trade over low-cost platforms. As discussed in Section 5.5, our evidence for adverse selection may also reflect the emergence of HFT in FX before our sample period.

The *Info* coefficients for real-money funds, private clients, MNCs, and SMEs are statistically insignificant, indicating that their spreads are not influenced by client information. This group includes 1,504 clients, 62% of the total, though they account for just 5% of trade value. This option is highlighted as potentially rational in Naik et al. (1999) but has not previously been documented empirically, to our knowledge. The no-response strategy is rational for clients that are uninformed or whose trades generate roughly equal adverse-selection costs and information-chasing benefits, and the evidence in Table 5.4 suggests that these clients are simply uninformed. A lack of information among non-financial firms is logical because they primarily use FX as a medium of exchange and internal governance mandates treat it as a cost within the finance function. Indeed, the bylaws of many non-financial firms discourage speculative trading because it brings rogue-trader risk, which is costly to manage (Osler, 2006).

The lack of information in the trades of real-money funds and private clients is not consistent with standard assumptions in economic theory. It is, however consistent with the observation of industry experts that real-money funds and private clients tend to treat currency trading as an unavoidable administrative cost rather than a source of returns (Taylor and Farstrup, 2006; Galanek, 2010). The incentive of such funds to invest in information may also be limited by the infrequency of their trades (Table 5.2). Bjønnes et al. (2021) similarly conclude that low-leverage asset managers and non-financial clients tend to be uninformed. Our evidence strengthens that finding because we measure client information directly rather than indirectly.

Theory does not dictate a specific time horizon for calculating post-trade returns when measuring client information, so we experiment with an alternative, 30 minutes, a period that is quite long given the speed and market penetration of HFT (Chaboud et al., 2014). This adjustment leaves most of our qualitative findings unchanged (Table 5.5, column 2). The information coefficients that were previously significant – those for hedge funds, CTFs, client banks, and brokers – retain their original signs but are smaller, presumably because the standard deviations for *Info* increase with time horizon. The *Info* coefficient for client banks becomes insignificant, suggesting that their trades only carry information about high-frequency returns. The estimated effects on half-spreads of a one-standard-deviation rise in *Info* remain higher for CTFs, at (now) 1.04 pip than for hedge funds, at 0.78 pip. The estimated effect of a one-standard-deviation decline in broker *Info* falls to 0.29 pip. *Info* remains insignificant for real-money funds, private clients, MNCs, and SMEs.

In summary, the evidence presented in this section confirms and expands Ramadorai's (2008) evidence for vast differences between order-driven and quote-driven markets with respect to asymmetric information across trade counterparties. In order-driven markets, such information asymmetries generate adverse selection that raises the equilibrium spread by the same amount for all traders (Glosten and Milgrom, 1985). In OTC markets with rational dealers, asymmetric information between dealer and client can raise the spread, lower the spread, or leave it unchanged, and the magnitude of any effect varies with the client's information.

5.4.1 Other Model Implications

The model's remaining coefficients are consistent with their motivating theories and the adjusted R^2 is encouragingly high at 0.58. Though one cannot directly compare explanatory power across datasets and models, it is nonetheless notable that explanatory power in previous studies of OTC client execution costs has been far lower, with a maximum of 0.33 (Hau et al., 2021). We review the coefficient estimates only briefly because they are not central to our analysis; a thorough discussion with economic impacts is available in a companion piece, Bjønnes et al. (2025).

All three of the remaining client characteristics generate significant dealer responses. More active clients benefit from narrower markups, as predicted in Bernhardt et al. (2005), as do more sophisticated clients, as predicted in Duffie et al. (2005) and Green et al. (2007). Consistent with quote skewing, markups are positively related to the predictability of a client's trade direction and the effect diminishes with sophistication. Consistent with evidence in Hendershott et al. (2020) and Hau et al. (2021), markups are larger (smaller) for clients with the fewest (most) alternative trading relationships.

Markups are also influenced by order-handling costs and intertemporal smoothing; inventory effects appear to be weak. The extra order-handling costs of direct trades are indeed passed on to clients; direct trading raises the markup by 2.7 pips on an average-sized trade. Markups on all trades are inversely related to trade size, consistent with fixed costs. Intertemporal smoothing is revealed by coefficients that would otherwise violate standard theory: dealers move markups inversely with volatility and with the daily pattern of core spreads. The core spread is insignificant. These coefficients might appear to violate inventory effects (Ho and Stoll, 1981), but each one is the sum of an inventory effect and another effect of opposite sign. We infer that inventory effects are relatively weak.

5.4.2 Extensions and Robustness Tests

Our results prove robust to five extensions or modifications of our baseline model (Table 5.6).

Brokers. We examine whether the dealer treats retail forex brokers and institutional brokers differently by including them separately. The results indicate that dealers seek the information of both client types (Table 5.6, Column 1) but the *Info* coefficient for institutional brokers, at -2.35, is ten times that for retail brokers, -0.22. This difference is logical because institutional brokers are often the agents of hedge funds while forex retail traders are generally uninformed (Heimer and Simon, 2014). The model's other qualitative implications remain unchanged.

Measuring market volatility. Our conclusion that markups are inversely related to volatility is robust to using realized volatility instead of the high-low range over the previous hour, HL_t. Realized volatility is calculated with one-minute Reuters mid-quote returns over the same pre-trade hour (Table 5.6, Column 2). The model's other qualitative implications also remain unchanged.

Measuring sophistication. We replace the SBP and MRFQ platform shares with an indicator variable for clients who are known to be expert in FX trading, specifically hedge funds and non-LFD client banks. The model's qualitative conclusions remain unchanged though the estimated dealer responses to client information increase in magnitude for hedge funds and brokers (Table 5.6, Column 3). With these alternative coefficients, the annual cost of a 1-standard-deviation rise in client information would be 3.3 percentage points for the average hedge fund and the annual savings would be 25.4 percentage points for the average broker. The coefficient on *PB* declines, which is logical because all *PB* clients are hedge funds. All other coefficients remain consistent with theory.

Market activity. We limit the sample to European and New York business hours, 8 – 21 GMT, when EUR-USD trading is most active. The *Info* coefficients again strengthen for hedge funds and brokers but the model's qualitative implications for the dealers' response to client information remain unchanged (Table 5.6, Column 4).

Price shading. We extend the model by adding dealer inventories, to test whether FX dealers shade client prices. Price shading has been documented for many OTC bond markets (e.g., Friewald and Nagler, 2024) but research has found no evidence for it in FX, as discussed in Section 5.2. Dealers have long confirmed that they do not shade client prices, citing two reasons: (i) the core market is highly liquid and inexpensive and thus a better alternative for managing inventory, and (ii) it can be costly to reveal their inventory position to clients such as hedge funds. We measure inventory as the dealers' net EUR accumulation over the most recent 100 EUR-USD trades (roughly 20 minutes, on average); prior research shows that active FX dealers return inventories to the desired level within five minutes of a shock (Bjønnes and Rime, 2005). This measure is not ideal, because the dealing floor's EUR inventory will also be influenced by trades against other currencies; however, those trades are not included in our data. Our inventory measure proves insignificant and its inclusion leaves the model's qualitative implications unchanged (Table 5.6, column 5).

CTF Information. We examine the robustness of our findings for CTFs by excluding their five largest trades, which had an extraordinarily large mean 1-minute return, 4.0 pips (Table 5.6, Column 6). Results are unchanged.

5.5 Information Chasing versus Adverse Selection

This section examines how dealers choose between information chasing and adverse selection when trading with informed clients. We provide evidence that the choice hinges in part on trade latency. Dealers widen spreads more aggressively with client information for better-informed HFT hedge funds than for other hedge funds.

The choice between narrowing and widening the markup for better-informed clients should be determined by the balance between adverse-selection costs and information-chasing benefits. After a trade with an informed client, the dealer will lose if the market moves adversely before inventory is restored (Glosten and Milgrom, 1985). However, the dealer might gain by quickly adjusting his/her quotes to others (Pintér et al., 2022) or by opening a parallel proprietary position in the core market (Naik et al., 1999).

Our dataset's 253 trades with negative markups appear to capture information chasing in action and thus merit close scrutiny. As shown in Table 5.7, these trades tend to be both larger than other trades, with mean size €2.4, and better informed. Their post-trade returns are already positive at the one-minute horizon, though not significant; they achieve significance by the two-minute horizon with mean two-minute return of 0.47 pip. Two minutes is a long time in liquid 21st-century markets, long enough for dealers to adjust quotes to other clients or to create a parallel proprietary position in the core market. The remaining trades in our sample provide a striking contrast, with post-trade returns that are two orders of magnitude smaller and statistically insignificant.

Information chasing was standard operating procedure in FX during the 1980s and 1990s; indeed, the physical layout of a trading floor was designed to support this strategy, according to a former trading-desk manager (Clyde, 1996). The FX dealers' attitude towards informed clients shifted markedly, however, after high-frequency trading arrived in the mid-2000s (Markets Committee, 2011), for reasons that were not difficult to identify. One important HFT strategy, latency arbitrage, intensifies the dealer's picking-off risk with respect to stale quotes (Menkveld and Zoican, 2017). The HFT strategy of anticipating large trades as they move through the market (Hirshey, 2021) accelerates price discovery and thereby reduces the time available for dealers to exploit client information via parallel core-market trades. HFT traders also capture profits from riskless arbitrage opportunities – profits that were previously available only to dealers – and speeds the disappearance of those opportunities (Budish et al., 2015; Chaboud et al., 2014).

We suggest that liquidity provision for HFTs is relatively costly to dealers and provides smaller benefits, which has led FX dealers to shift from information chasing to adverse selection with respect to these HFT funds. To test this hypothesis, we partition hedge funds into those that do and do not engage in HFT, based on information provided in their websites, and re-estimate the model.[40] If the hypothesis is correct, the coefficient on client information will be larger for HFT hedge funds than other hedge funds. The results provide strong support for the HFT hypothesis (Table 5.8). The *Info* coefficient for HFT funds is large, at 1.08 pip, and over three times the *Info* coefficient for non-HFT hedge funds, 0.33. For a one-standard-deviation rise in hedge-fund information, the additional annual cost at the average trade frequency for hedge funds would be 3.2 percentage points more for an HFT hedge fund than a non-HFT hedge fund. This finding could explain the difference between the dealers' treatment of hedge funds and institutional brokers. Information in broker trades comes from their clients, rather than the brokers themselves, and those clients cannot engage in HFT due to inescapable lags in client-broker communication.

5.6 Concluding Remarks

This paper provides evidence for striking differences between quote- and order-driven markets in the channels through which asymmetric information influences bid-ask spreads. Using highly detailed data from the world's largest OTC market it makes two contributions to the literature.

First, the paper proposes and documents that quote-driven spreads are influenced by two distinct information asymmetries: (i) the asymmetry between dealer and client and (ii) the asymmetry among professional traders in the interdealer or core market. This conclusion emerges from the two-tier structure of OTC markets and the fact that client prices are set as a markup over the same-side core-market price. Using the Huang and Stoll (1997) model we provide evidence that adverse selection in the core market contributes 25% of the average client spread, and the fraction ranges from 35% to 74% for hedge funds, CTFs, client banks, and brokers, the clients that are most sensitive to execution costs.

[40] It proved infeasible to identify HFT trading among CTFs or client banks.

The paper's second contribution is to provide evidence that OTC dealers adopt three different strategies for responding to client information. We estimate an OTC-tailored model of the dealer's choice variable, the markup. The model comprehensively captures the spread determinants highlighted in the literature, including four client characteristics as well as intertemporal smoothing, a strategy introduced here. Results indicate that OTC dealers narrow spreads for better-informed brokers, consistent with information chasing (Naik et al., 1999; Pintér et al., 2022). OTC dealers widen spreads for better-informed hedge funds, commodity trading firms, and client banks, consistent with adverse selection (Glosten and Milgrom, 1985; Easley and O'Hara, 1987). Finally, dealers make no information-related adjustment to spreads for clients they consider uninformed, specifically low-leverage asset managers ("real-money funds"), private clients, and non-financial firms. Previous studies of OTC markets provide empirical support for information chasing (Ramadorai, 2008; Pintér et al., 2022) but not for adverse selection (e.g., Reiss and Werner, 1996; Bernhardt et al., 2005; Goldstein et al., 2007) or the decision to not respond.

We hypothesize that HFT is a critical determinant of the dealers' choice between information chasing and adverse selection for informed clients. To test this hypothesis we partition hedge funds into those that do and do not engage in HFT and re-estimate the markup decomposition model. Results indicate that dealers widen spreads more aggressively for better-informed HFT hedge funds than for non-HFT hedge funds, consistent with our hypothesis.

Future research could helpfully analyze the extent to which these conclusions hold in other OTC markets. It would be valuable to learn how dealers respond to client information in illiquid OTC markets without extensive HFT, such as those for US corporate or municipal bonds. It would also be valuable to learn the extent to which interdealer information asymmetries influence client spreads in other OTC markets.

Table 5.1: Descriptive Statistics for Core-market Trades

Table shows summary statistics for liquidity-providing order-driven core-market trades. A cluster of trades have the same date and time. Data include all EUR-USD trades by a top-10 EUR-USD dealing bank during the 68 trading days from 2 January through 20 April 2012.

	All	Clustered		
		All	Last look	Firm-quote
All trades ≥ €250K				
Number trades	101,820	88,376	27,936	60,440
Trade size (€ mn)				
Mean	1.134	1.307	0.898	1.496
Median	1.000	1.000	1.000	1.000
St. dev.	0.823	1.138	0.637	1.263
All trades				
Number trades	131,269	117,805	56,925	60,899
Trade size (€ mn)				
Mean	0.891	0.993	0.465	1.485
Median	1.000	1.000	0.200	1.000
St. dev.	0.854	1.127	0.618	1.263

Table 5.2: Client Trades, Descriptive Statistics

Table shows summary statistics for client markups measured in pips, where one pip is \$0.0001/€. For client *c*, $BuySell_c \equiv |\#Buys_c - Sells_c| / (\#Buys_c + \#Sales_c)$, $BuySell_c \in [0,1]$. Data include all market-making EUR-USD trades by a top-10 EUR-USD dealing bank with non-dealing-bank counterparties during the 68 trading days from 2 January through 20 April 2012.

	All	Hedge Fund	CTF	Client bank	Broker	Real-money	Private Clients	MNC	SME
A. Basic properties									
Number clients	2,388	43	19	731	91	257	88	169	990
Number trades	257,241	6,624	19,792	73,730	150,152	1,185	113	1,261	3,660
Mn. trades/client	108	154	1,042	101	1,658	5	1	7	4
Trade size (€ mn)									
Mean	0.49	1.20	0.58	0.88	0.25	2.08	0.17	2.53	0.23
St. dev.	*1.91*	*1.81*	*0.49*	*3.15*	*0.61*	*6.60*	*0.30*	*5.48*	*0.84*
B. Markups (pips)									
Mean	0.44	0.03	0.01	0.26	0.00	2.59	38.35	3.09	22.26
St. Dev.	*3.96*	*0.48*	*0.37*	*2.51*	*0.13*	*11.59*	*17.19*	*10.80*	*18.39*
C. Daily Trading (€ mn)									
Mean	1647.46	102.46	148.01	832.16	481.36	31.69	0.24	40.93	10.61
St. Dev.	*400.39*	*60.06*	*77.40*	*258.46*	*155.79*	*29.66*	*0.72*	*23.93*	*6.79*
D. Platform shares (% of trades)									
Direct	1.87	1.97	0.03	2.04	0.04	13.80	100.00	24.57	67.59
SBP	29.61	1.17	0.03	68.57	15.31	32.83	0.00	17.68	30.52
MRFQ	27.15	47.00	57.03	26.11	22.51	53.18	0.00	57.75	0.94
API	41.37	49.86	42.92	3.29	62.14	0.19	0.00	0.00	0.96
E. BuySell									
Mean	0.11	0.09	0.03	0.25	0.04	0.29	0.23	0.29	0.70
St. Dev.	*0.20*	*0.09*	*0.05*	*0.27*	*0.04*	*0.34*	*0.39*	*0.33*	*0.41*

Table 5.3: Effective Core Spreads

Table reports coefficients from estimating Equation (5.2), the decomposition model for core spreads:

$$\Delta P_t = \frac{S}{2}(D_t - D_{t-1}) + \alpha \frac{S}{2} D_{t-1} + \beta \frac{S}{2} \Delta I_t + e_t,\ e_t \equiv \eta_t + \varepsilon_t - \varepsilon_{t-1}.$$

P_t: traded price in pips from an order-driven core-market platform; D_t: signed trade direction, +1 (-1) when the counterparty buys (sells); I_t: dealer inventory; S: effective spread in pips; α: adverse selection contribution as share of effective spread; β: price shading as share of effective spread. Data comprise complete spot EUR-USD trading record by a top-10 dealing bank during the first 68 trading days of 2012. Trades at the same price, date, and time are clustered as one transaction. Estimation is by GMM, robust standard errors clustered by time and trading platform. Significance at 10%, 5%, 1% and 0.1% levels indicated by *, **, *** and †, respectively.

	All trades ≥ €250K	**Active hours only**	**All sizes**	**No price shading**
Platform Type	1	2	**3**	4
Firm-quote order-driven				
Effective half-spread (e-4), *S/2*	0.611†	0.553†	0.580†	0.575†
Adverse selection, α (share of *S/2*)	0.608†	0.616†	0.636†	0.576†
Price shading, β (share of *S/2*)	0.040†	0.035†	0.039†	
Last-look order-driven				
Effective half-spread (e-4), *S/2*	0.139†	0.137†	0.229†	0.133†
Adverse selection, α (share of *S/2*)	0.900†	0.874†	0.738†	0.892†
Price shading, β (share of *S/2*)	0.054*	0.050	0.039***	
Adj. R^2	0.036	0.033	0.011	0.035
N. Obs.	88,375	72,071	117,805	88,375

Table 5.4: Client Information

Table reports information measures by client type. Data include all market-making EUR-USD trades by a top-10 dealing bank with non-dealing-bank counterparties during the 68 trading days from 2 January through 20 April 2012.

	Hedge Fund	CTF	Client Bank	Broker	Real Money	Private Clients	MNC	SME
A. Post-trade Returns, 1 min. (pips)								
All trades	0.349†	-0.017	-0.013	0.056†	0.066	-0.381	0.046	0.054
St. Dev. by trade	*3.069*	*3.277*	*3.044*	*3.366*	*2.783*	*2.720*	*2.957*	*3.077*
St. Dev. by client	*0.95*	*0.80*	*1.01*	*0.74*	*1.33*	*2.564*	*1.54*	*2.505*
[€0,€0.5) mn	0.357†	-0.100***	-0.056†	0.041†	0.065	-0.375	-0.240	0.406
St. Dev.	*2.848*	*3.170*	*3.017*	*3.427*	*2.514*	2.775	*2.593*	*3.071*
[€0.5,€1.0) mn	0.430†	0.046*	0.068**	0.065***	-0.039	0.500	-0.256	0.552
St. Dev.	*3.178*	*3.380*	*3.173*	*2.991*	*2.961*	1.924	*2.855*	*3.197*
[€1.0,€5.0) mn	0.271†	0.042*	0.079**	0.194†	0.100	-2.333	0.069	-0.152
St. Dev.	*3.324*	*3.346*	*3.024*	*3.287*	*3.000*	0.577	*3.111*	*3.078*
[€5.0,€∞] mn	0.284*	4.000**	0.178**	0.220	0.233	–	-0.207	-0.367
St. Dev.	*2.896*	*4.528*	*0.813*	*3.264*	*2.655*	–	*2.940*	*2.781*
B. Post-trade Returns, 30 min. (pips)								
All trades	0.577†	-0.005	-0.011	-0.020	-0.234	0.602	-0.304	0.054
St. Dev.	*13.852*	*15.175*	*14.266*	*14.451*	*12.908*	*16.17*	*14.320*	*15.073*
C. Num. Obs.								
All trades	6,624	19,792	73,730	150,876	1,185	113	1,261	3,660
[€0,€0.5) mn	1,392	8.409	41,429	120,486	416	104	248	3,335
[€0.5,€1.0) mn	1,838	1.526	11,419	18,056	296	6	133	165
[€1.0,€5.0) mn	3,090	9,852	17,874	11,777	341	3	763	145
[€5.0,€∞] mn	304	5	1,038	557	132	–	117	15

Table 5.5: Client Information and Client Markups

Table reports coefficients from estimating Equation (5.3), the markup decomposition model:

$$Markup = \sum_j InfoCT'_j \pi_j + \beta TV + Pltshr\delta_2' + \theta Skew + PCP\psi'$$
$$\varphi_0 + Time_t \kappa_1' + (\lambda_1 + \varphi_1)Sz + DT\ (\varphi_2 + \varphi_3\ Sz) + (\kappa_2 + \lambda_2)\ HL + (\kappa_3 + \lambda_3)\ CSprd + \zeta,$$
$$j = \{HF, CTF, Client\ Bank, Broker, RM, PC, MNC, SME\}.$$

Markup: absolute difference between the client's price and the prevailing core price. *Info:* the client's average post-trade return. CT_j: indicator for client type, j = {*HF, CTF, ClientBank, Broker, RM, PC, MNC, SME* }. *TV*: client's log trading volume over the sample period. *Pltshr*: clients' share of trades handled directly, over SBPs, or over MRFQs. *Skew*: Product of *ShrDT*, the client's share of direct trades, and *BuySell*, the predictability of client *c*'s trade direction. *PCP*: indicators for four client types: any client with a prime broker, SMEs that do not trade over MRFQs, private clients, and client banks from countries with low financial development that do not trade over MRFQs. *Time*: indicators for European hours, New York hours, or overnight trading (Asian hours excluded). *Sz*: log of trade *t*'s absolute amount in euros. *DT*: indicator for direct trades. *HL*: (log) high-low range over the previous hour. *CSprd*: core spread. Data include all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012. Robust standard errors clustered by date. Significance at the 5%, 1%, and 0.1% levels indicated by **, ***, and †, respectively.

	1-min *Info*	**30-min *Info***
Client Information		
x Broker	**-0.512†**	**-0.138†**
x Hedge Fund	**0.519†**	**0.287†**
x CTF	**1.120†**	**0.456***
x Client Bank	**0.305*****	**0.004**
x Real-money	**0.358***	**-0.049**
x Private Client	**-1.196**	**-1.142**
x MNC	**-0.739**	**-0.156**
x SME	**0.320**	**0.046**
Controls		
Trading Volume	-0.134†	-0.129†
Client Sophistication		
Direct (e-2)	2.803†	2.860†
Skew (e-2)	4.833†	4.811†
SBP (e-2)	-0.127†	-0.131†
MRFQ (e-2)	-0.163†	-0.201†
Potential Counterparties		
Prime Broker	-0.369†	-0.405†
$ClBk_{LFD\text{-}NM}$	0.622†	0.613†
Priv. Client	23.651†	24.042†
SME_{NM}	10.762†	10.784†
A. Smoothing		
Time: Europe	0.122†	0.121†
NYC	0.105†	0.094†
Overnight	-0.044*	-0.046*
Volatility (e-2)	-0.070**	-0.087***
Core Spread (e-2)	0.665	0.689
Order-handling and Inventory Costs		
Trade Size (e-2)	-0.300	0.442*
Direct Trade	32.227†	32.321†
DT x TrdSz	-2.079†	-2.086†
Constant	2.242†	2.072†
Adj. R^2	0.580	0.580
N. Obs.	257,180	257,180

Table 5.6. Robustness Tests

Table reports coefficients from estimating Equation (5.3), the markup decomposition model:

$$Markup = \sum_j InfoCT'_j \pi_j + \beta TV + Pltshr\delta'_2 + \theta Skew + PCP\psi'$$
$$\varphi_0 + Time_t \kappa'_1 + (\lambda_1+\varphi_1)Sz + DT\ (\varphi_2+\varphi_3\ Sz) + (\kappa_2+\lambda_2)\ HL + (\kappa_3+\lambda_3)\ CSprd + \zeta,$$
$$j = \{HF, CTF, Client\ Bank, Broker, RM, PC, MNC, SME\}.$$

Variables in the baseline regression are defined in Table 5.5. *Brokers*: Brokers disaggregated into online retail platforms and institutional advisors. *Alternative volatility*: Realized volatility over the previous hour replaces the high-low range. *Alternative sophistication*: Indicator for clients that are FX experts (hedge funds and non-LFD client banks) replaces platform shares for SBPs and MRFQs. *8–21 GMT*: Include only the hours of most active trading. *Price Shading*: Add net inventory accumulation over previous 100 trades (about 20 minutes). *CTF*: Exclude the five large CTF trades. Data include all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012. Robust standard errors clustered by date. Significance at the 10%, 5%, 1%, and 0.1% levels indicated by *, **, ***, and †, respectively.

	Brokers	**Alternative Volatility**	**Alternative Sophist.**	**8-21 GMT**	**Price Shading**	**CTF**
	1	**2**	**3**	4	5	6
Client Information						
x Broker		**-0.512†**	**-0.834†**	**-0.624†**	**-0.512†**	**-0.512†**
x Retail forex	**-0.217†**					
x Institutional	**-2.346†**					
x Hedge Fund	**0.545†**	**0.513†**	**0.778†**	**0.910†**	**0.519†**	**0.519†**
x CTF	**0.998†**	**1.121†**	**1.179†**	**1.366†**	**1.121†**	**1.136†**
x Client Bank	**0.312*****	**0.306*****	**0.351*****	**0.349****	**0.305*****	**0.305*****
x Real-money	**0.358***	**0.357***	**0.349***	**0.407***	**0.358***	**0.358***
x Private Client	**-1.196**	**-1.196**	**-1.196**	**-0.787**	**-1.196**	**-1.196**
x MNC	**-0.740**	**-0.752**	**-0.722**	**-0.670**	**-0.740**	**-0.739**
x SME	**0.320**	**0.320**	**0.320**	**0.247**	**0.320**	**0.320**
Controls						
Activity Discount	-0.141†	-0.134	-0.175†	-0.166†	-0.134†	-0.134†
Client Sophistication						
Direct (e-2)	2.838†	2.801†	2.967†	3.684†	2.804†	2.804†
Skew (e-2)	4.839†	4.834†	4.805†	4.144†	4.832†	4.833†
SBP (e-2)	-0.065†	-0.127†	–	-0.153†	-0.127†	-0.127†
MRFQ (e-2)	-0.112†	-0.163†	–	-0.178†	-0.163†	-0.163†
Experts			-0.408†			
Potential Counterparties						
B. *Prime Broker*	-0.339†	-0.366†	-0.179†	-0.591†	-0.370†	-0.369†
$ClBk_{LFD-NM}$	0.612†	0.622†	0.355†	0.490†	0.621†	0.622†
Priv. Client	23.626†	23.643†	23.218†	22.838†	23.650†	23.651†
SME_{NM}	10.743†	10.756†	10.403†	10.644†	10.761†	10.762†
Smoothing						
Time: Europe	0.127†	0.123†	0.120†	–	0.124†	0.122†
NYC	0.106†	0.106†	0.052†	–	0.106†	0.105†
Overnight	-0.045*	-0.366†	-0.100†	–	-0.044*	-0.044*
Volatility (e-2)	-0.064*	-27.943**	-0.008**	-0.060	-0.070**	-0.070**
CSprd	0.007	0.007	0.007	0.008	0.007	0.007
Order-handling and Inventory Costs						
Trade Size (e-2)	0.310	-0.296	1.450†	-0.204	-0.297	-0.298
Direct Trade	32.252†	32.230†	31.965†	30.065†	32.228†	32.227†
DT x TrdSz	-2.081†	-2.079†	-2.067†	-1.860†	-2.079†	-2.079†
C. *Constant*	2.207†	2.241†	2.746†	2.854†	2.242†	2.242†
Adj. R^2	0.580	0.580	0.581	0.591	0.580	0.580
N. Observations	257,241	257,241	257,241	181,098	257,141	257,236

Table 5.7. Trades with Negative Markups

Data include all market-making client trades in EUR-USD by a top-10 dealing bank during the first 68 trading days of 2012. Significance at the 5% level indicated by **.

	Mean Size	**Client's Post-trade Return (pips)**		
		1 min	**2 min**	**5 min**
Markup < 0	2.40	0.217	0.472**	0.522
Other trades by clients with negative markups	2.08	0.117	0.096	0.201
Markup = 0	0.47	0.004	0.005	0.007
Markup > 0	0.67	-0.003	-0.002	0.000

Table 5.8: Adverse Selection vs. Information Chasing

Table reports coefficients from estimating Equation (5.3), the markup decomposition model:

$$Markup = \sum_j InfoCT'_j \pi_j + \beta TV + Pltshr\delta'_2 + \theta Skew + PCP\psi'$$

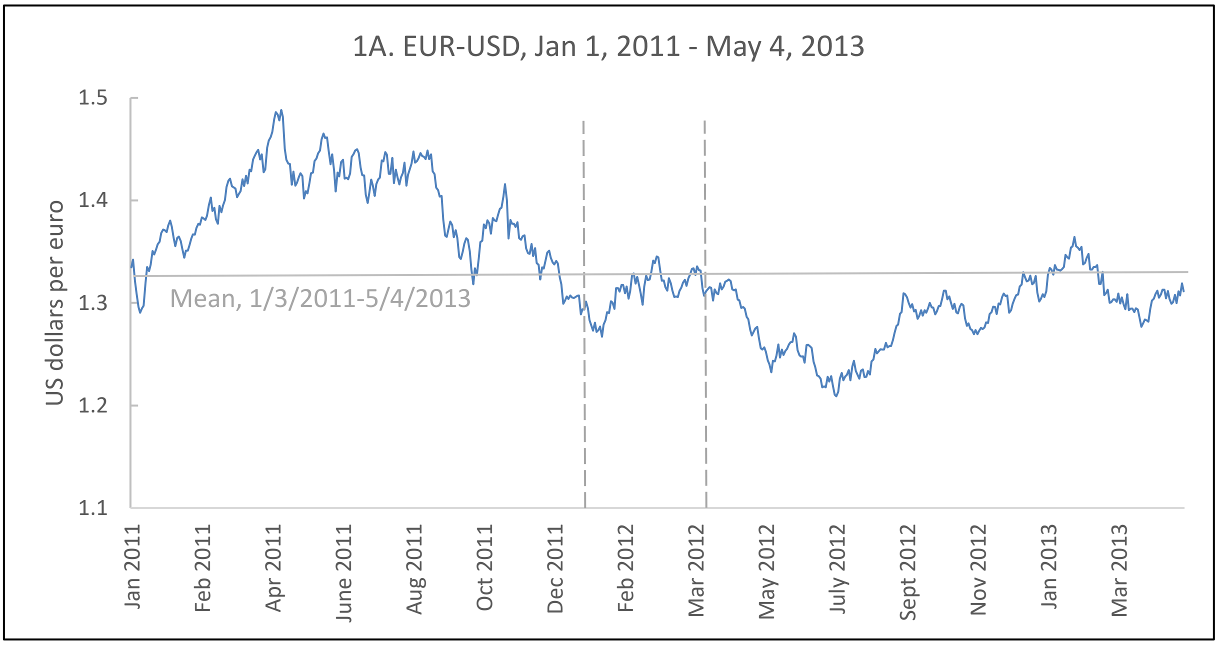
$$\varphi_0 + Time_t \kappa'_1 + (\lambda_1 + \varphi_1)Sz + DT\ (\varphi_2 + \varphi_3\ Sz) + (\kappa_2 + \lambda_2)\ HL + (\kappa_3 + \lambda_3)\ CSprd + \zeta\ ,$$

$$j = \{HF, CTF, Client\ Bank, Broker, RM, PC, MNC, SME\}.$$

Regression variables are defined in Table 4. Data include all market-making client trades by a top-10 EUR-USD dealing bank during the first 68 trading days of 2012. Robust standard errors clustered by date. Significance at the 5%, 1%, and 0.1% levels indicated by **, ***, and †, respectively.

	Info: 1-min.	**Info: 30 min.**
Client Information		
x Broker	**-0.514†**	-0.139†
x Hedge Fund		
Low Latency	**1.078†**	0.463†
Other	**0.333†**	0.220†
x CTF	**1.117†**	0.413*
x Client Bank	**0.304*** **	0.053
x Real-money	**0.357***	-0.045
x Private Client	**-1.196**	-0.142
x MNC	**-0.739**	-0.156
x SME	**0.320**	-0.146
Client Characteristics		
Trading Volume	-0.134†	-0.163†
Sophistication		
Direct (e-2)	2.816†	3.022†
Skew (e-2)	4.825†	4.778†
SBP (e-2)	-0.127†	0.000
MRFQ (e-2)	-0.160†	-0.160†
Access to Counterparties		
Prime Broker	-0.461†	-0.254†
$ClBk_{LFD\text{-}NM}$	0.622†	0.689†
Priv. Client	23.644†	23.635†
SME_{NM}	10.759†	10.442†
Order-handling and Inventory Costs		
Trade Size (e-2)	-0.258	2.115†
Direct Trade	32.215†	32.135†
DT x TrdSz	-2.079†	-2.078†
Constant	2.237†	2.496†
Smoothing		
Time: Europe	0.122†	0.117†
NYC	0.105†	0.048***
Overnight	-0.044*	-0.086***
Volatility (e-2)	-0.069**	-0.092**
Core Spread (e-2)	0.007	0.007
Adj. R^2	0.580	0.581
N. Obs.	257,241	257,241

Figure 5.1 EUR-USD Data sample realized volatility
EUR-USD level and realized volatility, 2 January 2011 through 4 May, 2013. Daily data. Realized volatility calculated from daily returns over centered 20-day periods.



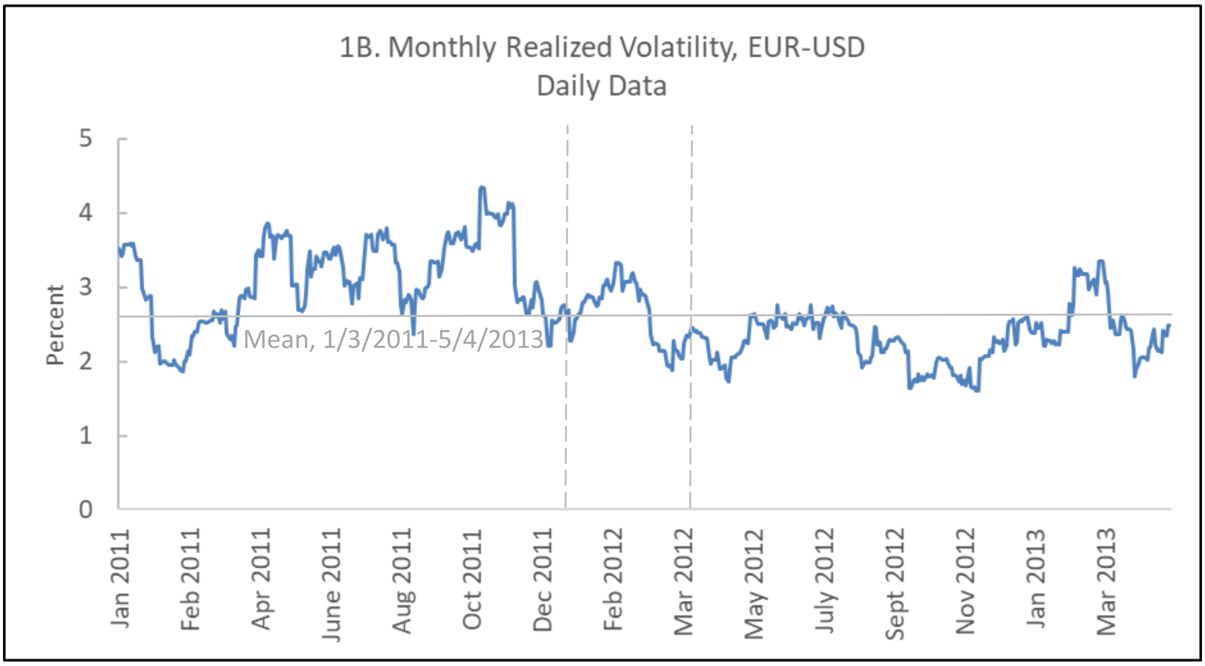
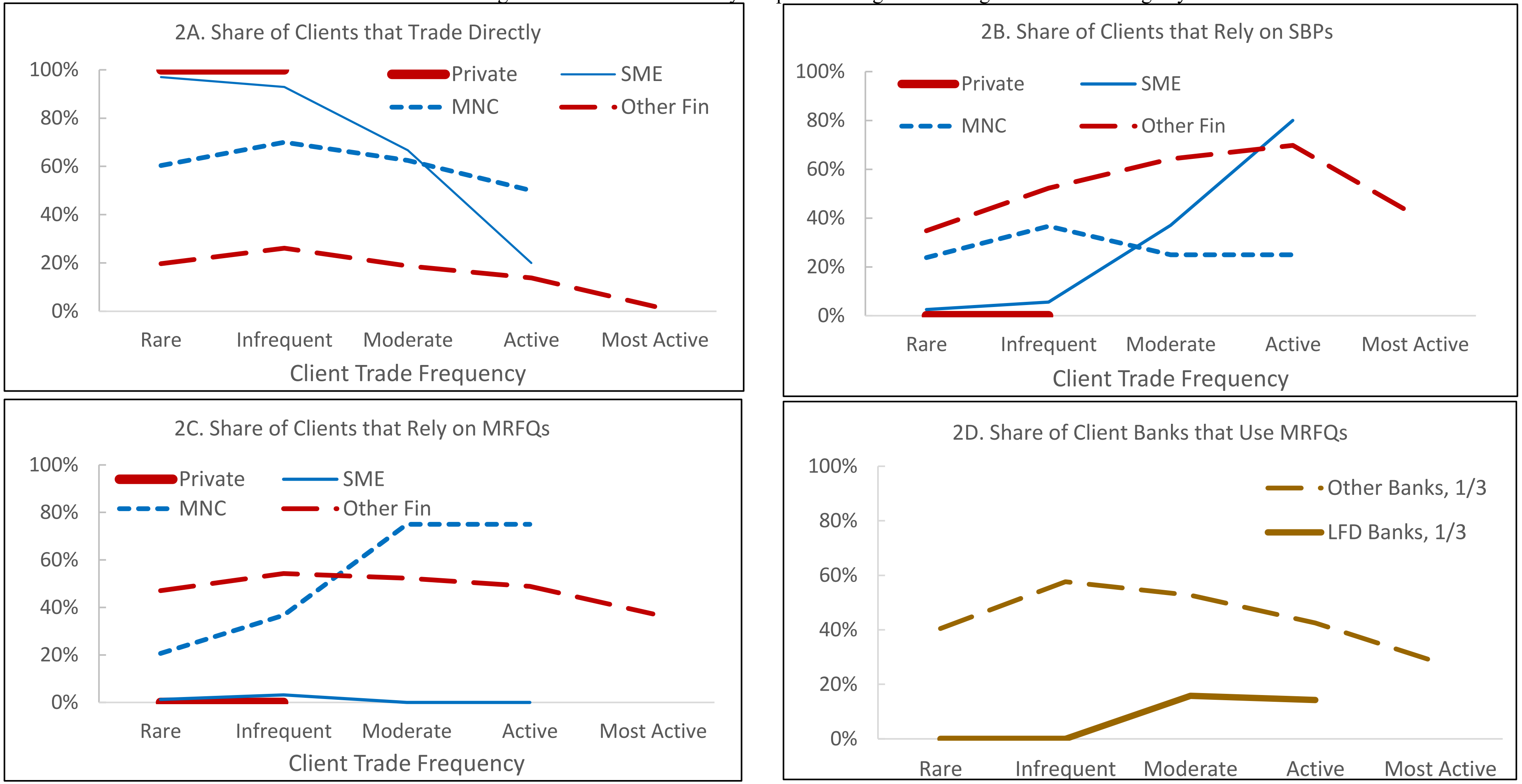


Figure 5.2 Client Sophistication and Institutional Constraints
Share of clients that rely on a given trading platform, by client trading frequency. *Rare* means no more than one trade per month. *Infrequent* means no more than one trade per week and more than *Rare*. *Moderate* means no more than one trade per day, and more than *Infrequent*. Active means no more than ten trades per day, and more than *Moderate*. *Most Active* means more than ten trades per day. *LFD client banks* come from countries in the lowest half of the IMF's Financial Development Index for 2012. *Other client banks* come from other countries. Data include all market-making EUR-USD client trades by a top-10 dealing bank during the first 68 trading days of 2012.



APPENDICES

6.1 Appendix to Chapter 3

Table 6.1.1 Robustness Tests based on Equation 3.2

	All time zones	Size threshold 0.25 mill.	No size threshold	Trade Size
Interbank OD	0.45	0.46	0.42	0.23
	[0.43, 0.47]	[0.43, 0.49]	[0.40, 0.43]	[0.21, 0.25]
Interbank RFQ	0.11	-0.03	0.08	0.03
	[0.09, 0.13]	[-0.05, -0.01]	[0.07, 0.09]	[0.01, 0.05]
Dark Pool	0.06	0.15	0.09	0.09
	[0.04, 0.08]	[0.12, 0.18]	[0.08, 0.10]	[0.07, 0.11]
MDP OD	0.17	0.09	0.13	0.15
	[0.16, 0.19]	[0.07, 0.11]	[0.12, 0.15]	[0.13, 0.17]
MDP RFQ	0.08	0.10	0.11	0.09
	[0.06, 0.09]	[0.08, 0.12]	[0.09, 0.12]	[0.06, 0.10]
SDP RFQ	-0.56	-0.23	-0.26	-0.01
	[-0.60, -0.52]	[-0.25, -0.21]	[-0.27, -0.25]	[-0.02, 0.01]

NOTE The table reports the long-run impulse responses of midpoint changes to unit innovations in trades, where trades are classified as buy (equal to 1) or sell (equal to -1). The corresponding 95% confidence intervals are included. Robustness checks are conducted using variations of Model 2 (as presented in Table 3). The first analysis incorporates all time zones, encompassing complete trading hours without restricting trades to European and U.S. trading sessions. The second SVAR specification lowers the trade size threshold included in the calculations to EUR250,000, while the third removes trade size limits entirely. The fourth robustness check incorporates a variation where order flow is measured as trade volume rather than as an indicator variable.

Table 6.1.2 Venue Classification

Venue	Example	Trading Protocol	Client Types
Interbank OD	EBS Markets, Reuters Matching	Limit Order Book	GLP, RLP and Hedge Funds
Interbank RFQ	Reuters Dealing, Bloomberg FXGO	Request for Quote	RLP, Client Banks
Dark Pool	BCG	Mid-match	GLP
Multi-Dealer Platform OD	Hotspot, Flextrade, Integral	Limit Order Book	Hedge Funds,GLP, RLP
Bilateral			
Multi-Dealer Platform RFQ	360T, Currenex, FX All	Request for Quote	Corporate clients, Client Banks
Single Dealer Platform*	Citi Velocity	Request for Quote	Brokers, Client Banks, Corporate clients

*includes RFQ via Electronic Brokers such as Tullet Prebon, Carl Kliem

NOTE The table classifies venue types in our dataset by execution types, customer types most commonly using the specific venue and representative industry venues for the respective category. Interbank order driven: Dealer banks trade with one another on electronic interdealer venues where best bid and ask are visible. Interbank direct (RFQ): Counterparty reaches out to an interbank dealer by a dealing system or telephone, email, txt, etc. Dark Pools, anonymous venues where transactions are executed on mid-quote prices. MDP order driven: Multi-Dealer Platforms with a central order book where clients can execute transactions against multiple dealing banks. MDP RFQ: Request for quote systems, one client requests near-simultaneous quotes from multiple dealing banks. SDP RFQ: Single-dealer platforms provide prices on the client's desktop and enables client to connect directly to the dealing bank. GLP: Global Liquidity Providers, ranked top 20 in Euromoney survey, RLP: Regional Liquidity Providers, ranked 21-50 in Euromoney survey, Client banks, smaller retail and commercial banks across the globe either ranked 51-100 in Euromoney survey or not ranked at all, Hedge Funds: Proprietary trading companies employing a combination of strategies: macro, technical and High-Frequency-Trading. Corporate clients: Medium size and Large Corporates active in the import and export business.

6.2 Appendix to Chapter 4

Numerical Solution Method for the general model

The general model is solved via sample average approximation (SAA). An overview of this method and its theoretical properties can be found in e.g. Homem-de Mello and Bayraksan (2014). The main idea of SAA methods is to turn a stochastic optimization into a deterministic one. This is achieved by first fixing a large sample S of the random variables involved and approximating theoretical moments such as expectation and variance by their sample counterparts. One then performs a numerical optimization routine on this fixed sample, and notes that the specific so lution to the deterministic problem converges to the solution of the stochastic problem as the sample size of S increases.

In the solution of the general model laid out in Section 4.2.2, we have set the parameters to $N = 30$ trading rounds, where the transaction time k is in round 15. The price impact parameter λ is set to 0.3 and the white noise-term in the pricing equation is Gaussian iid with unit variance. We use a sample size of 5000 draws of the random sequence (ϵ_n, $n = 1, 2, \ldots, N$).

BIBLIOGRAPHY

Abramowitz, M., & Stegun, I. A. (1964). *Handbook of mathematical functions: With formulas, graphs, and mathematical tables* (Vol. 55). Courier Corporation.

Anand, A., & Subrahmanyam, A. (2008). Information and the intermediary: Are market intermediaries informed traders in electronic markets? *Journal of Financial and Quantitative Analysis*, 43(1), 1–28.

Bank for International Settlements, 2020. The 2012 private sector involvement (PSI) in Greece. *Discussion Paper Series No. 11*.

Bank for International Settlements. (2022). *Triennial Central Bank Survey: Global foreign exchange market turnover in 2022*.

Barclay, M. J., Hendershott, T., & McCormick, D. T. (2003). Competition among trading venues: Information and trading on electronic communications networks. *Journal of Finance*, 58(6), 2637–2665.

Barclay, M., Hendershott, T., & Kotz, K. (2006). Automation versus intermediation: Evidence from Treasuries going off the run. *Journal of Finance*, 61(6), 2395–2414.

Barclay, M. J., & Warner, J. B. (1993). Stealth trading and volatility. *Journal of Financial Economics*, 34, 281–305.

Bech, M. L., Chapman, J., & Garratt, R. (2016). Hanging up the phone: Electronic trading in fixed income markets and its implications. *BIS Quarterly Review, March, 35–54*.

Berger, D. W., Chaboud, A. P., Chernenko, S. V., Howorka, E., Krishnasami Iyer, R. S., Liu, D., & Wright, J. H. (2008). Order flow and exchange rate dynamics in electronic brokerage system data. *Journal of International Economics, 75(1), 93–109*.

Bernhardt, D., Dvoracek, V., Hughson, E., & Werner, I. (2005). Why do larger orders receive discounts on the London Stock Exchange? *Review of Financial Studies,* 18, 1343–1368.

Bernhardt, D., & Taub, B. (2008). Front-running dynamics. *Journal of Economic Theory*, *138(1),* 288–296.

Bessembinder, H. (1994). Bid-ask spreads in the interbank foreign exchange markets. *Journal of Financial Economics, 35(3), 317–348*.

Bessembinder, H., & Venkataraman, K. (2004). Does an electronic stock exchange need an upstairs market? *Journal of Financial Economics, 73(1)*, 3–36.

Bessembinder, H., Maxwell, W., & Venkataraman, K. (2006). Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics, 82*, 251–288.

Biais, B., & Green, R. (2007) "The Microstructure of the Bond Market in the 20th Century", *IDEI Working Paper, n. 482*,

Bjønnes, G. H., & Rime, D. (2005). Dealer behavior and trading systems in foreign exchange markets. *Journal of Financial Economics, 75(3)*, 571–605.

Bjønnes, G. H., Osler, C. L., & Rime, D. (2021). Price discovery in two-tier markets. *International Journal of Finance & Economics, 26(2)*, 3109–3133.

Bjønnes, G. H., Kathitziotis, N., & Osler, C. (2025). Cross-client variation in OTC execution costs. Mimeo.

Bloomfield, R., & O'Hara, M. (1999). Market transparency: Who wins and who loses? *Review of Financial Studies, 12(1),* 5–35.

Bollerslev, Tim, and Michael Melvin, 1994. Bid-ask spreads and volatility in the foreign exchange market. *Journal of International Economics 36*: 355-372.

Budish, E. B., Cramton, P., & Shim, J. J. (2015, November). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics, 130*(4), 1547–1621. https://doi.org/10.1093/qje/qjv027

Cai, F. (2003). Was there front running during the LTCM crisis? Unpublished manuscript.

CBOE Global Markets. (2023a). U.S. equities market volume summary, https://www.cboe.com/us/equities/market_share/.

CBOE Global Markets. (2023b). European equities market share by market. https://www.cboe.com/europe/equities/market_share/market/venue/#dm=tbpcan&dr=5day&mt=1&ms=0&hc=1&f=0&ID=e6bf20a14833c25152c7&V=65d60609d1021c727921.

Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *Journal of Finance, 69*(5), 2045–2084.

Chaboud, A., Rime, D., & Sushko, V. (2023). *The foreign exchange market* (BIS Working Paper No. 1094). Bank for International Settlements. https://www.bis.org/publ/work1094.htm

Chakravarty, S., & Li, K. (2003). An examination of own account trading by dual traders in futures markets. Journal of Financial Economics, 69(2), 375–397.

Chaturvedula, C., Bang, N. P., Rastogi, N., & Kumar, S. (2015). Price manipulation, front running, and bulk trades: Evidence from India. *Emerging Markets Review, 23(C),* 26–45.

Cheung, Y.-W., & Chinn, M. D. (2001). Currency traders and exchange rate dynamics: A survey of the U.S. market. *Journal of International Money and Finance, 20(4),* 439–471.

Clyde, W. 1996, *Intraday foreign exchange market anomalies and microstructure*, working paper, Quinnipiac College, August

Collin-Dufresne, P., Junge, B., & Trolle, A. B. (2020). Market structure and transaction costs of index CDSs. *Journal of Finance, 75(5),* 2719–2763.

Collin-Dufresne, P., Hoffman, P., & Vogel, S. (2019. Informed traders and dealers in the FX forward market. https://sammf.com/wp-content/uploads/2020/10/Collin-Dufresne.pdf.

Collin-Dufresne, Pierre, and Vyacheslav Fos, 2015. Do prices reveal the presence of informed trading? *Journal of Finance 70(4):* 1555-1582.

Comerton-Forde, C., & Tang, K. M. (2007). Anonymity, frontrunning, and market integrity. *The Journal of Trading, 2(4),* 101–118.

Conlin, I. E. A. (2023). Electronic credit trading approaching inflection point in IG. *Tradeweb*. Retrieved April 5, 2023, from https://www.tradeweb.com/newsroom/media-center/insights/commentary/electronic-credit-trading-approaching-inflection-point-in-ig

Conrad, J., Johnson, K. M., & Wahal, S. (2003). Institutional trading and alternative trading systems. *Journal of Financial Economics, 70(1),* 99–134.

Copeland, T. E., & Galai, D. (1983). Information effects on the bid-ask spread. *Journal of Finance, 38(5),* 1457–1469.

Cornell, B. (1978). Determinants of the bid-ask spread on forward foreign exchange contracts under floating exchange rates. *Journal of International Business Studies*, 9(2), 33–41.

Danthine, J.-P., & Moresi, S. (1998). Front-running by mutual fund managers: A mixed bag. *Review of Finance, 2(1),* 29–56.

de Mello, T. H., & Bayraksan, G. (2014). Monte Carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science, 19(1),* 56–85.

Demsetz, H. (1968). The cost of transacting. *Quarterly Journal of Economics, 82(1),* 33–53.

Dhokakia, Utpal M., 2015. The risks of changing your prices too often. *Harvard Business Review (6 July).* https://hbr.org/2015/07/the-risks-of-changing-your-prices-too-often.

Duffie, D., Gârleanu, N., & Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica,* 73(6), 1815–1847.

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

Edwards, A., Harris, L., & Piwowar, M. (2007). Corporate bond market transaction costs and transparency. *Journal of Finance, 62(3),* 1421–1451.

Evans, M. D. D., & Lyons, R. K. (2002). Order flow and exchange rate dynamics. *Journal of Political Economy, 110(1)*, 170–180.

Evans, M. D. D., & Lyons, R. K. (2005). Meese-Rogoff redux: Micro-based exchange-rate forecasting. *American Economic Review:* Papers and Proceedings, 95(2), 405–414.

Evans, Martin, and Richard K. Lyons, 2012. Exchange rate fundamentals and order flow. *Quarterly Journal of Finance 2(4):* 1-63.

Feldhütter, P., & Poulsen, T. K. (2018). What determines bid-ask spreads in over-the-counter markets? (SSRN Electronic Journal). https://doi.org/10.2139/ssrn.3286557

Fishman, M. J., & Longstaff, F. A. (1992). Dual trading in futures markets. *The Journal of Finance,* 47(2), 643–671.

Froot, K. A., & Ramadorai, T. (2005). Currency Returns, Intrinsic Value, and Institutional-Investor Flows. *Journal of Finance, 60*(3), 1535–1566.

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

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

Glosten, L. R., & Harris, L. (1988). Estimating the components of the bid-ask spread. *Journal of Financial Economics, 21(1)*, 123–142.

Goldstein, M., & Hotchkiss, E. S. (2020). Providing liquidity in an illiquid market: Dealer behavior in US corporate bonds. *Journal of Financial Economics, 135(1),* 16–40.

Goldstein, M., Hotchkiss, E. S., & Sirri, E. R. (2007). Transparency and liquidity: A controlled experiment on corporate bonds. *Review of Financial Studies, 20(2),* 235–273.

Green, R. C., Hollifield, B., & Schurhoff, N. (2007). Financial intermediation and the costs of trading in an opaque market. *Review of Financial Studies, 20(2),* 275–314.

Hagströmer, B. (2021). Bias in the effective bid-ask spread. *Journal of Financial Economics,* 143(1), 314–337.

Harris, L. (2003). Trading and exchanges: Market microstructure for practitioners. Oxford University Press.

Harris, L., & Piwowar, M. S. (2006). Secondary trading costs in the municipal bond market. *Journal of Finance, 61(3),* 1361–1397.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance, 46(1)*, 179–207.

Hatheway, F., Kwan, A., & Zheng, H. (2017). An empirical analysis of market segmentation on U.S. equities markets. *Journal of Financial and Quantitative Analysis,52(6*), 2399–2427.

Hau, H., Hoffmann, P., Langfield, S., & Timmer, Y. (2021). Discriminatory pricing of over-the-counter derivatives. *Management Science, 67(11),* 6660–6677.

Hau, H., Killeen, W., & Moore, M. (2002). How has the euro changed the foreign exchange market? *Economic Policy, 17(34),* 149–192.

Hendershott, T., & Madhavan, A. (2015). Click or call? Auction versus search in the over-the-counter market. *Journal of Finance, 70(1),* 419–447.

Hendershott, T., Li, D., Livdan, D., & Schurhoff, N. (2020). Relationship trading in OTC markets. *Journal of Finance, 75(2),* 683–734.

Hirschey, N. (2021). Do high-frequency traders anticipate buying and selling pressure? *Management Science, 67*(6), 3321–3345. https://doi.org/10.1287/mnsc.2020.3608

Ho, T. S. Y., & Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics, 9(1),* 47–73.

Hollifield, B., Neklyudov, A., & Spatt, C. (2017). Bid-ask spreads, trading networks, and the pricing of securitizations: 144A vs. registered securitizations. *Review of Financial Studies, 30(10),* 3048–3085.

Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies, 10(4),* 995–1034.

Ito, T., & Hashimoto, Y. (2006). Intraday seasonality in activities of the foreign exchange markets: Evidence from the electronic broking system. *Journal of the Japanese and International Economies, 20(4)*, 637–664.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77(6),* 1121–1134.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica, 53(6),* 1315–1335.

Lagos, R., & Rocheteau, G. (2009). Liquidity in asset markets with search frictions. *Econometrica, 77(2)*, 403–426.

Love, R., & Payne, R. (2008). Macroeconomic news, order flows, and exchange rates. *Journal of Financial and Quantitative Analysis, 43(2),* 467–488.

Lyons, R. K. (1995). Tests of microstructural hypotheses in the foreign exchange market. *Journal of Financial Economics, 74(2),* 321–351.

Madhavan, A. N., & Smidt, S. (1993). An analysis of changes in specialist inventories and quotations. *Journal of Finance, 48(5)*, 1595–1628.

Markets Committee, (2011). *High-frequency trading in the foreign exchange market*. Basel: Bank for International Settlements.

Markham, J. W. (1988). Front-running—Insider trading under the Commodity Exchange Act. *Catholic University Law Review, 38*, 69–90.

Mende, A., Menkhoff, L., & Osler, C. L. (2006). Asymmetric information and the cross-section of currency spreads. *Journal of Financial Markets, 9(3),* 223–245.

Menkhoff, L., & Schmeling, M. (2010). Whose trades convey information? Evidence from a cross-section of traders. *Journal of Financial Markets, 13(1),* 101–128.

Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2016). Information flows in foreign exchange markets: Dissecting customer currency trades. *Journal of Finance, 71(2),* 601–634.

Menkveld, A. J., & Zoican, M. A. (2017). *Need for speed? Exchange latency and liquidity. The Review of Financial Studies, 30*(4), 1188–1228.

Moore, M., & Payne, R. (2011). On the sources of private information in FX markets. *Journal of Banking & Finance, 35(5)*, 1250–1262.

MSRB (Municipal Securities Rulemaking Board), 2018. How are municipal bonds quoted and priced? https://www.msrb.org/sites/default/files/how-are-municipal-bonds-quoted-and-priced.pdf

Naik, N. Y., Neuberger, A., & Viswanathan, S. (1999). Trade disclosure regulation in markets with negotiated trades. *Review of Financial Studies, 12(4),* 873–900.

Newey, W. K., & McFadden, D. (1994). Large sample estimation and hypothesis testing. In R. F. Engle & D. L. McFadden (Eds.), *Handbook of Econometrics (*Vol. 4, pp. 2111–2245). Elsevier.

O'Hara, M. (2003). Liquidity and price discovery. *Journal of Finance, 58(4),* 1335–1354.

O'Hara, M., & Zhou, X. A. (2021). The electronic evolution of corporate bond dealers. *Journal of Financial Economics, 140(2)*, 368–390.

O'Hara, M., Wang, Y., & Zhou, X. A. (2018). The execution quality of corporate bonds. *Journal of Financial Economics, 130(2),* 308–326.

Osler, Carol, 2006. Macro lessons from microstructure. *International Journal of Finance and Economics 11*: 55-80.

Osler, C. (2009). Foreign exchange microstructure: A survey. In *Encyclopedia of Complexity and System Science* (pp. 3682–3706). Springer.

Osler, C., Mende, A., & Menkhoff, L. (2011). Price discovery in currency markets. *Journal of International Money and Finance, 30(2)*, 1696–1718.

Osler, C., & Savaser, T. (2022). Shrouding and the FX trades of global custody banks. *Journal of Banking & Finance, 136*, 106393.

Payne, R. (2003). Informed trade in spot foreign exchange markets: An empirical investigation. *Journal of International Economics, 61(2),* 307–329.

Peterson, M. A., & Sirri, E. R. (2003). Order preferencing and market quality on U.S. equity exchanges. *Review of Financial Studies, 16(2),* 385–415.

Phillips, A., Stewart, A., & Adams, S. (2020). *TCA and fair execution: The metrics that the FX industry must use* (LMAX Exchange White Paper). LMAX Exchange. https://www.lmax.com/documents/LMAXExchange-FX-TCA-Transaction-Cost-Analysis-Whitepaper.pdf

Pintér, G., Wang, C., & Zou, J. (2022). Information chasing versus adverse selection. Bank of England Staff Working Paper No. 971.

Pirrong, C. A. (2015). *The economics of commodity trading firms* (Trafigura white paper; abridged version). Trafigura. https://www.trafigura.com/media/k4ocz3zq/2015_trafigura_economics_of_commodity_trading_firms_abridged_en.pdf

Ramadorai, T. (2008). What determines transaction costs in foreign exchange markets? *International Journal of Finance & Economics,* 13(1), 14–25.

Ranaldo, A., & Santucci de Magistris, P. (2022). Liquidity in the global currency market. *Journal of Financial Economics, 146(3),* 859–883.

Ranaldo, A., & Somogyi, F. (2021). Asymmetric information risk in FX markets. *Journal of Financial Economics, 140(2)*, 391–411.

Reiss, P. C., & Werner, I. (1996). Transaction costs in dealer markets: Evidence from the London Stock Exchange. In A. Saunders & I. Walter (Eds.), *The industrial organization and regulation of the securities industry* (pp. 125–175). University of Chicago Press.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance, 39(4),* 1127–1139.

Saakvitne, J. (2018). Trading strategies and information flow around price benchmarks [Technical report]. BI Norwegian Business School.

Sannikov, Y., & Skrzypacz, A. (2014). Dynamic trading: Price inertia, front-running and relationship banking [Working paper].

Scopino, G. (2014). The (questionable) legality of high-speed ‘pinging’ and ‘front-running’ in the futures markets. Unpublished manuscript.

Securities Industry and Financial Markets Association. (2023a). *US corporate bond statistics*. Retrieved June 20, 2025, from https://www.sifma.org/resources/research/us-bond-market-statistics/

Securities Industry and Financial Markets Association. (2023b). *US municipal bond statistics*. Retrieved June 20, 2025, from https://www.sifma.org/resources/research/us-municipal-bond-statistics/

Stoll, H. R. (1978). The supply of dealer services in securities markets. *Journal of Finance, 33(4)*, 1133–1151.

Taylor, A., & Farstrup, A. (2006). *Active currency management: Arguments, considerations, and performance for institutional investors* (CRA RogersCasey International Equity Research Report). Darien, CT: CRA RogersCasey.

Tkatch, I., & Kandel, E. (2008). *Demand for the immediacy of execution: Time is money* [Working paper]. Hebrew University of Jerusalem. https://doi.org/10.2139/ssrn.685944

Van Kervel, V., & Menkveld, A. J. (2019). High-Frequency Trading around Large Institutional Orders. *Journal of Finance, 74(3),* 1091–1137.

Van Ness, B. F., Van Ness, R. A., & Warr, R. S. (2001). How well do adverse-selection components measure adverse selection? *Financial Management, 30*(3), 77–98. https://doi.org/10.2307/3666304

Yueshen, B. Z., & Zou, J. (2022). Less is more. SSRN Working Paper. SSRN. https://doi.org/10.2139/ssrn.4274063

Zhu, H. (2014). Do dark pools harm price discovery? *Review of Financial Studies, 27(3),* 747–789.

DECLARATIONS

8.1 Erklärung

Hiermit erkläre ich, Neophytos Kathitziotis, dass ich keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die Arbeit wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

_________________________ ______________________________

Ort/Datum Unterschrift Doktorand/in

8.2 Eidesstattliche Versicherung

Ich, Neophytos Kathitziotis, versichere an Eides statt, dass ich die Dissertation mit dem Titel:

„Essays in Market Microstructure“ selbst und bei einer Zusammenarbeit mit anderen Wissenschaftlerinnen oder Wissenschaftlern gemäß den beigefügten Darlegungen nach § 6 Abs. 3 der Promotionsordnung der Fakultät für Wirtschafts- und Sozialwissenschaften vom 18. Januar 2017 verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht benutzt.

_________________________ ____________________________________

Ort/Datum Unterschrift Doktorand/in

8.3 Selbstdeklaration

This cumulative dissertation consists of four research papers.

- Chapter 2 was co-authored with Carol Osler and Geir Bjønnes. The research idea, methodological design, and manuscript were developed jointly. I was responsible for building and cleansing the dataset, implementing the econometric analysis. Overall, my contribution amounted to 50% in concept and planning, execution, and manuscript preparation.
- Chapter 3 was conducted entirely independently. I formulated the research question, designed the methodology, collected and analysed the data, and wrote the manuscript. My contribution for this paper is 100%.
- Chapter 4 was co-authored with Jo Saakvitne and Geir Bjonnes during my research visit to BI Business School. The idea and methods were jointly developed, and the data analysis and manuscript writing were shared equally, with my contribution amounting to 50% across all stages.
- Chapter 5 was largely written while I was visiting Brandeis University. The paper idea, literature survey, and interpretation of findings were developed collaboratively with Carol Osler and Geir Bjønnes, while I independently built the database, selected and implemented the econometric techniques, and coded the empirical analysis. Overall, my contribution is 50% in concept and planning, execution, and writing.

For all co-authored papers, these percentage contributions were agreed upon with the co-authors.

________________________ __________________________________

Ort/Datum Unterschrift Doktorand/in