Title
The Microstructure of Financial Markets: Insights from Alternative Data

Author
Skiera, Vincent

Publication Date
2021

Peer reviewed|Thesis/dissertation

The Microstructure of Financial Markets: Insights from Alternative Data

by

Vincent Skiera

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Dmitry Livdan, Chair
Professor Terrence Hendershott
Professor Christine Parlour
Professor Robert Bartlett III

Fall 2021

The Microstructure of Financial Markets: Insights from Alternative Data

Abstract

The Microstructure of Financial Markets: Insights from Alternative Data

by

Vincent Skiera

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Dmitry Livdan, Chair

This dissertation aims to promote our understanding of financial markets, particularly focusing on over-the-counter (OTC) markets. Contrary to exchanges, where all participants can trade at the same prices, in OTC markets, trading is bilateral. On exchanges trading is anonymous, so participants do not know with whom they are trading. In OTC markets bilateral trading implies that each participant knows with whom they are currently trading. Therefore, intermediaries, called dealers or liquidity providers, provide personalized prices to each participant. Understanding the provision of personalized prices is an important area of research. In the past, data limitations provided obstacles for this area of research.

Before the introduction of Trade Reporting and Compliance Engine (TRACE) in 2002, data in OTC markets were generally not available. TRACE collects and timely disseminates all transactions involving dealers in the corporate bond market, making information on corporate bond prices and traded quantities available to all participants, not just the participants of the trade. While TRACE is specific to the US corporate bond market, similar Reporting Facilities emerged in the municipal bond market. Once introduced, TRACE revealed that transaction costs, the difference between the price paid and the value of the corporate bond, were high and very variable across clients. Over time, TRACE has revealed additional information, most notably a unique identifier for each dealer. This identifier allows us to study how prices, and thus transaction costs, vary across different dealers.

To this day, TRACE or the corresponding repositories of all trades in other markets, like the municipal bond market, are common data for analyzing OTC markets. The data includes all trades and identifies the dealers in the transaction. However, many OTC markets do not have a repository collecting all trades, most notably the foreign exchange (FX) markets, an OTC market, and the world's largest financial market. Also, the existing repositories have multiple drawbacks.

Firstly, repositories do not reveal the identity of a client in a trade. This lack of identification means that only outcomes for the average client of a dealer can be studied. One cannot observe differences across clients. Thus differences across dealers for different clients cannot be determined. Furthermore, from a client's point of view, differences across dealers are only comparable if the client is the average client for each dealer, i.e., the average client across dealers is the same.

Secondly, repositories do not observe trade initiation. So it cannot be determined which participant contacted the other participant asking to trade. Instead, the literature commonly assumes that clients always initiate trades when trading with dealers. However, for interdealer trades, trade initiation cannot be assigned to either dealer in this way. Furthermore, in riskless principal transactions, where a dealer trades with two clients simultaneously, only one client initiated the trade with the dealer, while the dealer likely initiated the trade with the other client. However, which of the two clients initiated the trade can also not be assigned.

Thirdly and finally, the repositories only collect information on trades. Many participants have multiple dealers from whom they elicit quotes when wanting to trade. Repositories do not observe any of the quotes that dealers provided but did not lead to trade. However, quotes reveal important information about trading and the market functioning. Quotes reveal which dealers clients contact to trade and whether these dealers respond. In addition, quotes allow to measure competition in trades. More generally, quotes allow to observe in much greater detail why transaction cost differ across clients.

To counter these shortcomings, research has to rely on alternative data. This dissertation utilizes data from a leading trading platform in the FX market that addresses these shortcomings. The Bank of International Settlement's "Triennial Central Bank Survey of Foreign Exchange and Over-the-counter (OTC) Derivatives Markets in 2019" mentions six multi-bank platforms in the FX market. One of those six platforms provides the data for this dissertation. The data allows to identify both dealers and clients and provides information about the trade initiator, i.e., the liquidity demander. Most importantly, however, the data observes the quotes that liquidity demanders receive from all their dealers. To my knowledge, it is the first research to observe and utilize these three pieces of information.

In the first chapter **Chapter 1: Client-Dealer Intermediation in OTC Markets**, I utilize said data to analyze the prices that small clients receive in OTC markets and how they can improve their prices. I show that clients realize better prices when contacting more dealers, and they receive better prices from a particular dealer when trading more with the dealer. Small clients can only satisfy one of these conditions, as trading with many dealers means trading only little with each dealer. Utilizing the quote data, I find that clients have a third option. They can trade with a type of dealer I call "match maker." Upon being contacted by a client, a match maker immediately contacts other dealers, receiving quotes from those dealers and relaying the best quote she receives to her client at a mark-up. The match maker then trades with the dealers if and only if the client trades with her. Only alternative data, in this case, quote data, can unveil this type of dealer, as I need to observe the behavior of these dealers when they do not trade.

Match makers effectively pool their clients trading together under one identity and thus overcome the disadvantages of being small. In doing so, match makers enlargen their client's networks of dealers by at least 40% and reduce the trading cost for their small clients by 20%. I show that match makers are prevalent, with 14% of dealers being match makers and 12% of clients trading with them.

In the second chapter **Chapter 2: Centrality in OTC Markets, Liquidity Provision, and Prices**, I study the question of how the choice and number of dealers affect the prices that clients receive. This choice is a much-studied question in many different OTC markets, generally finding that more central participants profit when trading, i.e., a centrality premium exists. Central participants are participants with many trading partners, while peripheral participants have few trading partners. This research is the first to observe trade initiation, i.e., liquidity provision. As in the previous literature, without conditioning on liquidity provision, I find a centrality premium, i.e., the more central participant makes a profit in the trade. However, the more central participant generally provides liquidity in a trade. Once I control for liquidity provision, I find that all liquidity demanders, independent of their centrality, pay a spread for demanding liquidity.

Furthermore, this spread is larger the less central the liquidity demander is. The centrality of the liquidity demander mainly determines the size of the spread, the centrality of the liquidity provider is less important. Central liquidity demanders receive quotes from many liquidity providers simultaneously. For them, I find that more central participants trade more often with them. At the same time, the trade prices are similar across the centrality of the liquidity provider. Peripheral liquidity demanders only contact one or a few liquidity providers. For them, I find that they trade at better prices when trading with more central liquidity providers. Taken together, I infer for all liquidity demanders that the more central a liquidity provider is, the better the average price that the liquidity provider quotes. Thus, I find a centrality discount once controlling for trade initialization.

The centrality premium observed in previous research is due to liquidity provision by more central participants. Conditioning on the liquidity provision in a trade, I find a centrality discount.

The first two chapters of this dissertation study how transaction costs differ across clients in the FX market, with a focus on especially small clients. In particular, the first chapter shows how small clients can reduce their trading costs. In both chapters small clients are small corporations or small institutional investors, still trading at least hundred of thousands a month. Truly small investors are households. Compared to institutional investors, many households find it difficult to access financial markets. The final chapter explores whether a particular form of financial innovation, namely Robo-Advisers, help making financial markets more accessible to households.

In the final chapter **Chapter 3: Robo-Advisers: Household Stock Market Participation and Investment Behavior**, I study whether households benefit from financial innovation. Like small clients in the OTC market, households find it difficult to participate in financial markets. However, instead of facing high monetary costs, many households lack the proper education, i.e., financial literacy, to be willing to participate in financial markets. Financial literacy is generally low and difficult to improve. In this work, I study whether financial innovation, in the form of a Robo-Adviser, encourages households to invest in financial markets. This work is the first to study a fully automated investment process that offers advice to households with much lower net worth than human financial advisers require. I find that one-third of all Robo-Adviser users are new to financial markets, significantly higher than for regular retail investors. Furthermore, these households would not have invested in financial markets were it not for the introduction of the Robo-Adviser. I conclude that Robo-Advisers are an important tool to encourage participation in financial markets, providing households with academically-vetted portfolios.

This dissertation is dedicated to my parents and my wife Geunah, with love and deep gratitude for their unwavering support, encouragement, and care.

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I sincerely thank my committee chair: Professor Dmitry Livdan. He provided me exceptional guidance and very constructive criticism, and light-hearted discussion about everyday life, including our joint hobby of cycling. I am immensely grateful for his mentorship during the program. His exceptional guidance ensured my academic and physical well-being.

Next, I thank Professor Terrence Hendershott, Professor Christine Parlour, and Professor Joel Hasbrouck. They extraordinary generously provided me much of their time, such that my dissertation has benefited substantially from their extremely valuable suggestions and feedback.

Furthermore, I thank my data provider who have asked to remain anonymous. They showed faith in me and gave me an incredible, unique opportunity to cooperate with them and provide me with this dissertation's data. They showed patience during the lengthy research process and were always there to help me understand financial markets better. Without them, this dissertation would not be what it is.

So many other people deserve thanks and gratitude for helping me on this journey. For helpful comments and suggestions on my research, I am very grateful to Robert Bartlett III, Tobias Berg, Thomas Ernst, Brett Green, Albert "Pete" Kyle, and Ulrike Malmendier. For helpful discussions about financial markets and the connection to my research, I thank Tim Cook, Sebastian Hofmann-Werther, Simon Jones, Colm Murtagh, Martin Schmidberger, Milen Stanchev, Ralf Stromeyer, Moritz Weber, and Davina Zell. Sergio Conti, Corinne Faure, Holger Kraft, Joachim Schleich, Thomas Stephan and Karl-Theodor Sturm supported me to pursue a PhD in the first place. I thank Bo Jiang, who gave me the certainty that I continue my journey along the right path for me.

I am grateful for research support in Chapter 3 of my thesis by the Think Forward Initiative.

The Ph.D. program is a long journey; the bonds formed with fellow students made this program so much more meaningful. I have truly enjoyed discussing ideas, sharing academic and life advice, and goofing around with Sam Agarwal, Konhee Chang, Sylvain Delalay, Marius Guenzel, Maris Jensen, Dominik Jurek, Jaeyeon Lee, Mohammad Rezaei, Chaoran Yu, Ao Wang, Fan Zhang, Yipei Zhang, Miles Zheng and Harry Zhou, as well as many others.

I thank my parents for always picking up the phone and showing their love and care. My mother teaching me the grit and perseverance needed for a Ph.D. and her constant motherly care for my well-being. My father, who was also my role model, continued to serve as a very special mentor, providing encouragement and guidance.

Finally and most importantly, I sincerely thank Geunah Jang, whom I am proud to call my wife. Her unwavering love, no matter how far away, and full faith in me was the greatest source of encouragement.

Chapter 1

Client-Dealer Intermediation in OTC Markets

1.1 Introduction

In over-the-counter (OTC) markets, smaller clients receive worse prices than larger clients, with scholars, market participants, and regulators expressing concerns about the accessibility of these markets. Hundreds and sometimes thousands of dealers provide prices in OTC markets, so numerous opportunities to trade exist. However, these concerns appear warranted since it seems that small clients are discriminated against because they are small. Clients with whom a dealer trades higher volumes trade at better prices with this dealer (Bernhardt, Dvoracek, Hughson and Werner, 2004; Edwards, Harris and Piwowar, 2007; Di Maggio, Kermani and Song, 2017). OTC markets allow for price discrimination, as trading is non-anonymous and clients rely on repeated interaction with a small subset of the hundreds of dealers. The current research seeks to provide a detailed characterization of why small clients are price discriminated against, in order to explore means by which small clients mitigate the price disadvantages they experience.

To achieve these goals, I analyze quote-level data from a leading multi-bank platform in the FX market. Crucially, I show that small clients can reduce their transaction costs by pooling together and appearing as one "larger" client. Moreover, my data enable me to document that many dealers in OTC markets provide such a pooling service, which indeed reduces clients' transaction costs. Though my data suggest that such dealers are prevalent and commonly used by clients, to my knowledge, the current study is the first to explicitly document the nature of the service that these dealers provide.

In particular, my analysis reveals that Client A receives a worse price from a dealer than Client B does if and only if Client A trades less with this dealer than Client B does. This finding holds independent of the respective clients' sizes. I confirm this finding using both the simultaneous provision of quotes to many clients by a single dealer and the spread between the bid and ask prices that the dealer offers different clients.

My data further reveal that, compared with larger clients, smaller clients not only trade with fewer dealers but also receive quotes from fewer dealers. This distinction is not driven by a tendency on dealers' part to respond less frequently to smaller clients' requests for quotes (RFQs); indeed, my data reveal that dealers generally respond to all client requests, regardless of client size. Rather, the tendency of smaller clients to receive fewer quotes compared with larger clients is driven by smaller clients' tendency to request quotes from fewer dealers—despite the fact that contacting more dealers, all else equal, has the potential to reduce trading costs by enabling the client to select a dealer who is more willing to trade. Taken together, these findings explain why small clients tend to experience higher transaction costs compared with larger clients: They trade at smaller volumes with each of their dealers, and contact very few dealers for quotes.

The nature of this price discrimination implies that smaller clients could reduce their transaction costs by pooling with other clients to appear under one "identity". When the one "identity" approaches a particular dealer, it creates the impression of larger size, i.e., higher trade volume, eliciting better prices than they would obtain by contacting the same dealer independently.

Indeed, my data show that such a pooling service exists in the form of dealers that exclusively specialize in matching clients to other dealers (client-dealer intermediation) via riskless principal transactions (RPT). I refer to these dealers as "match makers". Match makers are numerous in my data (I identify 46 match makers, 14% of the dealers in my data set). They operate as follows: Whenever a client contacts a match maker in an RFQ, the match maker immediately initiates her own RFQ, contacting her dealers, and providing the client her best quote at a markup. Only if the client decides to trade with the match maker does the match maker execute the same trade with her dealers. When the match maker initiates her own RFQ, her dealers do not observe her client, but observe instead only the match maker. Thus, in effect, the match maker combines all of her clients under a single "identity".

Another advantage that the match maker provides her clients, beyond creating a shared identity in submitting RFQs to other dealers, is in potentially expanding those clients' network. Though the match maker does not possess a unique advantage over other clients in establishing her own network—my results show that a match maker requests quotes from the same number of dealers that any similar-sized client would approach—she might nevertheless request quotes from more, or different, dealers than her clients would. My analysis confirms that this is, in fact, the case, such that the match maker increases the number of dealers that directly or indirectly provide quotes to each client (indirect network size).

Specifically, the match maker provides her average client with indirect access to 6.4 more dealers than the client accesses directly. This (indirect) increase in the scope of the client's dealer network is similar across clients of different sizes: The match maker provides small and large clients with indirect access to quotes from 8 and 6 additional dealers, respectively (with the difference being attributable to the fact that large clients contact more dealers, such that there is an overlap of 2 dealers between their networks and the match maker's network). Even clients who exclusively request quotes from a single match maker obtain access to 1.5 more dealers compared with other clients of similar size who do not trade with match makers.

My analysis shows that, in pooling clients together and expanding their indirect networks, match makers attain prices for their clients that are 0.4 bps better than the prices those clients would have attained from a similar-sized "regular" dealer (i.e., a dealer that holds inventory rather than serving exclusively as an intermediary between the client and another dealer). In other words, the match maker reduces her clients' transaction cost by 20%, even considering the fee that the match maker charges for the RPT.

It is important to acknowledge that, while match makers use RPT for 99% of their trading volume, small trades making up 1% of their trading volume are not passed on as RPT and instead are taken into inventory[1]. Trades that are taken into inventory by the match maker are

[1]These trades are not immediately passed on to other dealers, and instead are generally passed on only later. The match maker collects these small trades that are not passed on (median trade size 14,948 EUR) and passes them

significantly more expensive than when traded with a similar-sized regular dealer. Therefore, the transaction cost advantages provided by match makers are limited to cases in which they use RPT; match makers are not good at providing additional liquidity to their clients by providing their own quotes.

All match makers are peripheral dealers in OTC markets, which are structured as core-periphery networks (Di Maggio et al., 2017; Hollifield, Neklyudov and Spatt, 2017; Li and Schürhoff, 2019; Hasbrouck and Levich, 2021). To the best of my knowledge, in revealing the existence of match makers, my work is the first to show that many peripheral dealers provide services that fundamentally differ from those of core dealers. I further observe that match makers are popular with clients, with 12% of clients trading with match makers. The capacity of match makers to obtain better prices for their clients by pooling them together and expanding their networks reveals, for the first time, a mechanism explaining why clients choose to trade with peripheral dealers.

In addition to formally introducing the concept of a match maker in OTC markets, the current study characterizes the clients who trade with such dealers and the manner in which they do so. A regular dealer executes less than 15% of her trading volume with clients that trade exclusively with the dealer. A match maker, however, executes 65% of her trading volume with exclusive clients, clients that only request quotes from her. Furthermore, most of these exclusive clients are small; in other words, match makers' pooling services are especially attractive to small clients. Match makers thus do not just behave differently from regular dealers after trading with clients, but also cater to different populations of clients.

The remainder of the paper is structured as follows. The following section reviews current literature on price discrimination in OTC markets, as well as on network structures in such markets. Section 1.3 describes the data. In Section 1.4, I show how transaction costs differ across clients, especially in relation to client size, and how clients choose their networks. Section 1.5 documents the existence and prevalence of match makers among dealers. Section 1.6 uses the results from Section 1.4 and related literature to derive possible reasons for clients to trade with match makers. Section 1.6 goes on to test these possible reasons and, in particular, to examine whether match makers indeed reduce their clients' transaction costs (relative to trading directly with regular dealers). Section 1.7 concludes.

1.2 Literature Review

Previous research on OTC markets has explored whether and how dealers price-discriminate between clients (Edwards et al., 2007; Green, Hollifield and Schürhoff, 2007). Studies in this vein (Bernhardt et al., 2004; Di Maggio et al., 2017; Di Maggio, Franzoni, Kermani and Sommavilla, 2019a; Li and Schürhoff, 2019; Hendershott, Li, Livdan and Schürhoff, 2020a) have shown

on to another dealer once their total is larger (e.g. above 100,000 EUR). The match maker thus takes hold of these trades for only a small amount of time and only for a small quantity.

that price discrimination can be attributed to the individual relationships between clients and dealers. In particular, several studies have documented that dealers provide better prices to clients that they trade more with, arguing that clients have greater bargaining power when trading more with a dealer (Bernhardt et al., 2004; Di Maggio et al., 2017, 2019a). However, why this greater bargaining power arises is unknown.

Indeed, given the opaque nature of OTC markets, much of the research on differences in trading behavior in such markets is limited by a lack of data. Until very recently (about 2017), most studies in corporate bond markets were unable to identify the trading parties (Edwards et al., 2007; Green et al., 2007; Zitzewitz, 2010; Harris, 2015). More recent work has identified dealers but not clients (Goldstein and Hotchkiss, 2020; Li and Schürhoff, 2019). Studies that identify both dealers and clients rely on non-public, proprietary data covering the market only partially (see, e.g., Hendershott et al. (2020a), who used data from the National Association of Insurance Commissioners). Yet a far greater limitation to the study of price discrimination is that most previous studies only observed trades but not quotes. While the data of Hendershott and Madhavan (2015) and Riggs, Onur, Reiffen and Zhu (2020) observe quotes at least to a partial extent, those studies do not study the differences between clients or between dealers.

Furthermore, as public datasets in OTC markets only identify trades and provide no further information about those trades, studies of such datasets tend to assume that most trades are the result of Nash bargaining (Duffie, Gârleanu and Pedersen, 2005; Üslü, 2019). However, the rise of electronic trading (O'Hara and Zhou, 2021) has brought about new trading mechanisms. Electronic platforms enable traders to easily contact multiple dealers simultaneously, as opposed to having to telephone each prospective dealer individually (Vogel, 2019). The first trading mechanism to emerge on electronic platforms was an auction mechanism, namely, the RFQ; this mechanism is at the focus of the current work. Prior works examining RFQs include Hendershott and Madhavan (2015), who explored when RFQs are used, and Riggs et al. (2020), who studied competition in RFQs. However, many aspects of the mechanisms of competition in RFQs are unknown. The current research leverages a quote-level and dealer and client identifying dataset to shed light on some of these open questions. In particular, my dataset enables me to explore how the relationship between a dealer and a client interacts with the competition in RFQ to affect pricing.

My paper further complements the literature on endogenous network formation in financial markets, and its implications for trading and prices. Due to limited availability of client identifiers in such markets, most previous work has focused on network choice among dealers (Di Maggio et al., 2017; Hollifield et al., 2017). For example, Upper and Worms (2004), Cocco, Gomes and Martins (2009) and Afonso and Lagos (2015) explored the structure of the interdealer market and the effect of this market structure on the prices that are negotiated in the market. Only a few studies have explored client networks (examples include Hendershott et al. (2020a) and Di Maggio, Egan and Franzoni (2019b)), so the characteristics of such networks are still largely unknown. Furthermore, previous studies have tended to assume that all dealers in a market perform similar functions. One exception is the work of Di Maggio et al. (2019b), who explored heterogeneity in non-trading dealer services; still, that study assumed that dealers show similar trading behavior. The current study observes the structures of client-dealer networks, and reveals, for the first time, a class of dealers—namely, match makers, mentioned briefly in the introduction and characterized in detail in what follows—whose trading behavior is fundamentally distinct from that of other types of dealers.

In identifying this distinction among dealers, my work further contributes to research on the positioning of dealers in the broader network structure of OTC markets. In particular, OTC markets are characterized by a core-periphery network structure (Li and Schürhoff (2019); Farboodi, Jarosch and Shimer (2017); Wang (2017), among others), in which a small number of dealers in the core conduct most trades, whereas a much larger number of peripheral dealers trade much less. Core dealers intermediate between clients and are willing to hold inventory between a client selling and a client buying. Models of OTC markets (Duffie et al. (2005); Wang (2017); Farboodi et al. (2017), among others) assume that all dealers intermediate between clients and hold inventory[2]; however, it is not known whether peripheral dealers indeed function in this manner. Moreover, research on interdealer trading has shown that dealers that are more peripheral receive worse prices when trading with core dealers (Di Maggio et al., 2017; Hasbrouck and Levich, 2021), putting peripheral dealers at an disadvantage when it comes to trading with clients. Likewise, in client-dealer trading, core dealers seem to provide lower prices than peripheral dealers do (Li and Schürhoff, 2019)[3], raising the question of why clients trade with peripheral dealers at all. In revealing the unique function of match makers—all of whom are peripheral dealers—my study contributes towards answering this question.

Given that I identify match makers on the basis of their use of RPTs—transactions in which the dealer, upon trading with a client, immediately passes the trade on to another dealer or client—my analysis provides a more in-depth understanding of use of RPTs in OTC markets. The common occurrence of RPTs in such markets was first discovered by Zitzewitz (2010) and later Harris (2015). Both of these studies focused mainly on retail trading, but showed that RPTs are common in institutional trades as well (Harris, 2015). Saar, Sun, Yang and Zhu (2019) used a theoretical model to explore the welfare effects of RPT with two clients, but RPTs with one dealer and one client are not understood. My analysis explains why clients are engaging in this type of RPT.

Finally, this work relates to the microstructure of FX markets (see Evans and Rime (2019) for a recent literature review). Most papers in this literature focus on interdealer trading in FX markets, as the continuous central limit order book is ideal for the study of price discovery (Evans and Lyons, 2002; Moore, Schrimpf and Sushko, 2016), and Hagströmer and Menkveld (2019) show that information diffuses from the central limit order book to the private quotes of

[2]Dealers may have different ability to perform this function or choose to differentiate themselves in how they serve their function, but in the end they all perform the same function.

[3]Li and Schürhoff (2019) show that shorter chains relate to lower prices and that more central dealers are involved in shorter chains. Even though a peripheral dealer may charge a less of a spread to trade, the longer intermediation chains still can lead to larger costs to the client.

dealers. Relatively few studies have explored the dealer-to-client market, as the very fragmented trading environment makes it difficult to obtain records with multiple dealers (Menkhoff, Sarno, Schmeling and Schrimpf, 2016). Yet, despite these difficulties, one is ultimately interested in client outcomes, as dealers exist so that clients can trade with each other. Similar problems plague the study of trading costs for different market participants, even in interdealer markets. However, Hasbrouck and Levich (2019) and Hasbrouck and Levich (2021), using CLS settlement data, identified settlement banks, and were the first to study the network structure of the FX market. The current study similarly overcomes data limitations and sheds light on dealer-to-client networks and relationships and trading costs in the FX market.

1.3 Foreign Exchange Market Structure and Data Description

1.3.1 Foreign Exchange Market Structure

The FX (spot) market is a heavily fragmented OTC market with fragmentation across and within market segments. The market comprises two segments: the interdealer market and the dealer-to-client market. The interdealer market covers trading between dealers and includes centralized options; in this market, trading takes place electronically and uses one of two limit order books (Hasbrouck and Levich, 2021): EBS (e.g., (Hagströmer and Menkveld, 2019)) and Reuters, besides bilateral trading. Each currency pair predominantly trades in a single limit order book, with EBS focused on EUR, JPY, and CHF; and Reuters being the dominant trading platform for GBP, AUD, CAD, and the Scandinavian currencies (King, Osler and Rime, 2013).

The interdealer market, which covers 30% of trading (Bank of International Settlement, 2019), plays a key role in price discovery (Moore et al., 2016; Hagströmer and Menkveld, 2019). In particular, Hagströmer and Menkveld (2019) found that the interdealer market (EBS in their study) first incorporates price updates, which later lead to updates in the prices quoted by dealers. At the same time, Evans and Lyons (2002) showed that prices in the FX market are heavily influenced by client order flow, finding R^2 of above 60%, so that price discovery is very much about the demand for currencies by clients. While these limit order books are accessible almost exclusively to dealers, a minority of interdealer trading takes place in limit order books, with a much greater share taking place in OTC markets (Schrimpf and Sushko, 2019).

As clients generally do not have access to the interdealer market, their trading occurs in the dealer-to-client market; as this is an OTC market, trading is non-anonymous. In contrast to many other OTC markets, trading in the dealer-to-client FX market is mainly electronic (71% of spot volume) (Bank of International Settlement, 2019). The dealer-to-client market is much more fragmented compared with the interdealer market, and is also much less studied.

The dominant form of client-dealer trading is through single- and multi-bank platforms (Bjønnes and Kathiziotis, 2016; Bank of International Settlement, 2019). These are electronic platforms where dealers either respond to clients' RFQs or stream prices to their clients. The difference between single-bank and multi-bank platforms is that, in single-bank platforms, only one dealer (and maybe the dealer's prime brokerage clients) provides prices and thus liquidity to clients. In contrast, in multi-bank platforms, multiple dealers provide prices to clients. Bjønnes and Kathiziotis (2016) show that dealers are active on both types of platforms and that multi-bank platforms are the more common form of trading, executing 39% of all trading volume versus 12% on single-bank platforms. My data set covers trading from a leading multi-bank platform.

Compared to many other OTC markets, the FX market is very liquid, and trades more volume than any other financial market, with more than $1 tr daily trading volume in the spot market. Similarly, the set of currencies is very limited, so that receiving a quote is not a problem[4]. The small number of currency pairs and ample trading of all currency pairs means that clients have no difficulty trading reasonable trade sizes[5], which my work focuses on. Therefore, I do not expect RFQs to be pre-arranged between clients and dealers; rather, a client is expected to contact a dealer only once the client wants to trade. Given that trades on the platform are conducted electronically, dealers should respond instantaneously.

Unreported analysis shows that Streaming is much less available than RFQ: While more than 1700 clients received quotes via RFQ, only 206 clients trade via Streaming. Moreover, a median client in RFQ was only larger than 10.7% of clients in Streaming, whereas the median Streaming client was larger than 95.5% of clients trading via RFQ. As this paper focuses particularly on small clients, I will focus on RFQ.

1.3.2 Trading on a Multi-Bank Platform

Trading on a multi-bank platform is non-anonymous, allowing dealers to personalize pricing, i.e., the provision of quotes, to each client. Before submitting a request, for a specified volume and currency pair, the client decides which dealers to contact and can contact any dealer that the client has a relationship with. After selecting the dealers to contact and determining the trade size and currency pair, the client can decide either to request both a sell and a buy price or to request a trade quote only in the direction of the desired trade. The default option is the latter, so that the vast majority of RFQs only receive quotes in the client's trade direction.

After receiving an RFQ from a client, the dealer decides whether to submit a quote and, if so, what quote to submit. Due to the fast-changing nature of the FX market and the tight pricing, dealers have the option to change their price up to the time of execution by the client or cancellation of the RFQ. A client selects at most one quote to trade, after which the dealer whose quote was selected has the right to reject the trade ("last look"). Though popular in other trading mechanisms, last look is uncommon in RFQ, as the dealers first observe the request by the client before providing quotes to the client.

[4]In my data only 0.5% of RFQs fail to receive a quote from a dealer; instead, most RFQs receive not one quote but many (6 on average).

[5]Round lots in FX markets are trades of 1 million, and in my data the 99th percentile is around 11.6 million.

1.3.3 Data Description

My data were provided by a leading multi-bank platform in the FX market, and focus on dealer-to-client trading. The platform is identified as one of only six multi-bank platforms in the Bank of International Settlement's discussion of the "Triennial Central Bank Survey of Foreign Exchange and Over-the-counter (OTC) Derivatives Markets in 2019" (Bank of International Settlement, 2019) in Schrimpf and Sushko (2019). Further anecdotal evidence supporting the platform's leading position in the market is reflected in the fact that, according to my discussions with platform managers, Citi—the third-largest dealer in the FX market (Euromoney, 2019)—continued to provide prices to this platform even after deciding, in late 2019, to reduce the number of platforms it provides pricing to by two-thirds (Financial Times, 2019).

As is typical in the FX market, trading on the platform is fully electronic. The dataset covers trading between 1,705 clients and 334 dealers in October 2019 (October 1st, 2019 to October 31st, 2019) via RFQ. The dealers and clients in the dataset form 9,453 dealer-client relationships, in which clients submitted 190,873 RFQs, each of which, on average, received quotes from 6 dealers. Within these 9,453 dealer-client relationships, 5,624 client-dealer pairs execute a trade, for a total of 117,176 executed RFQs in 178 currency pairs. For each RFQ, the data provides information on trade size and currency, identifying the client and the dealers asked for a quote and shows the last quote submitted by each dealer, thus showing the choice of quotes, i.e., trade prices, the client had when executing the trade. Further details about the sample are given in Appendix A.1.

In general, many dealers in FX markets utilize multiple trading venues to trade; for example, a large dealer might have its own single-bank platform and provide quotes on the given multi-dealer platform in addition to access to the interdealer market. With regard to the clients, however, my discussions with the data provider[6] give me reason to believe that most non-bank clients (1286 non-bank clients out of 1,705 total clients) exclusively trade on this multi-bank platform. Even among bank clients, I show in Skiera (2021a) that 97% of bank clients that request and provide liquidity on the platform, thus acting both as a client and as a dealer, request as least as much trade volume as they provide, indicating that they likely exclusively use this platform.

The fragmented nature of the dealer-to-client market has meant that no complete trading record of FX trading exists, with previous studies often relying on data from one dealer (Evans and Lyons, 2002; Menkhoff et al., 2016). More recently, progress has been made by Hagströmer and Menkveld (2019), who observed non-tradeable quoted prices across multiple dealers, and used these data to measure the information flow from the interdealer market to dealers. Likewise, Hasbrouck and Levich (2019) and Hasbrouck and Levich (2021) utilized settlement records to construct the network structure of the FX market, capturing interdealer trading

[6]Besides the opportunity to trade, the multi-bank platform providing the data also provides the client with many additional services, such as documentation on regulatory compliance, such as Mifid II, pre- and post-trade analytics, as well as integration into enterprise management systems.

even outside the limit order books and dealer-to-client trading across different settlement members. While reliance on data from a single trading platform does not cover all trading, a central repository covering all trades, which does not even exist in the FX market, also has it disadvantages. Central repositories in OTC markets do not report the identity of trading clients (Di Maggio et al., 2017; Li and Schürhoff, 2019) and often not the identity of dealers, e.g., in the Corporate Bond Market TRACE pre-2017 (Edwards et al., 2007). Furthermore, central repositories only identify trades and not provided quotes. Accordingly, to my knowledge, my paper is the first to utilize both client and dealer identification in trading; to observe all dealers whom a client approached with an RFQ; to observe all quotes provided by dealers to a client; and to observe the response of a client to the provided quotes (trade or no trade)—providing an unprecedented opportunity to study client networks and client trading behavior.

1.3.4 Descriptive Statistics

Table 1.1 reports descriptive statistics for the clients in the sample during the data collection period (October 2019). Panel A of Table 1.1 shows what types of clients are present in the market. The focus of this dataset on the dealer-client segment is reflected in the fact that the overwhelming majority of clients are non-banks (75%), with a strong focus on corporations. These non-bank clients execute almost all RFQs that they initiate (by volume). Only banks deviate from this behavior, initiating many more RFQs than they execute. (As I show below, a large part of this difference between the requested volume and executed volume is due to match makers, who initiate an RFQ whenever a client contacts them and execute that RFQ if and only if the client trades with them.) The "network size" of the average client—i.e., the number of dealers that provide the client with quotes in response to an RFQ (see below for further discussion of this definition)—is significantly larger than one would observe focusing only on trades, as the average client receives prices from 5.5 dealers, but trades with only 60% of them, showing that most relationships lead to infrequent trading.

Panel B of Table 1.1 shows a breakdown of trade volume across clients of different sizes (where a client's "size" is defined as its trade volume over the data collection period). Despite the fact that most relationships result in infrequent trading, I observe that trading among clients is very concentrated. Specifically, 10 (0.6%) clients trade 20% of all trading volume, and just over 11% of clients execute 80% of the trading volume. To execute this trading volume, clients contact many dealers, with each of the 190,873 RFQs receiving, on average, quotes from 6 dealers, as noted above. The descriptive statistics, however, show that these are mainly the same dealers for each client's request. For example, the largest clients, despite trading over 1bn, each receive quotes from only 27.1 dealers on average, showing that they form stable networks, as in Hendershott et al. (2020a). The smaller the client, the smaller the number of dealers from which that client receives a quote. As I will show in what follows, the differences across clients in the numbers of dealers that provide quotes (client network size) are not driven by the fact that that dealers respond less to smaller clients (e.g., Client group *31-75* receives a response from only 83% of dealers they request a quote from, while client group *1001+* receives a quote from 85% of dealers), but instead by the fact that smaller clients choose to contact fewer dealers. While the largest 10 clients each contact 28.3 dealers on average, the smallest 718 clients each contact only 3.51 dealers on average.

Panel C shows that the distribution of trade volume across dealers is similarly concentrated. In particular, the Top 5 dealers execute more than 20% of all trade volume; moreover, a Top 5 dealer provides quotes to more than 15% of clients, whereas the smallest 134 dealers provide quotes to only 1/100th the number of clients a Top 5 dealer has. The core-periphery structure of the FX market is further highlighted by the fact that Top 5 dealers each provide quotes in more than 20% of all RFQs (by count) and provide more than 500 times as much quoted volume than the smallest 134 dealers provide on average. The competition in the FX market is reflected in the low percentage of quotes that are converted into trades: All dealers convert less than one-tenth of the volume that they provide in quotes in response to RFQs[7].

1.4 Transaction Costs of Clients

1.4.1 Client Networks

The literature on OTC markets shows that transaction costs vary greatly across trades (Edwards et al., 2007), and auction theory tells us that a key determinant of seller revenue, in my case client transaction cost, is the number of bidders, in my case dealers, involved in the transaction, in my case an RFQ. Therefore, the size of a client's network (i.e., the number of dealers who provide the client with quotes in response to RFQs) is potentially a key factor in explaining the transaction costs a client faces. In order to test this proposition, a different number of dealers need to be contacted in different RFQ. Therefore, I study the heterogeneity of client networks.

I start by showing the distribution of client network sizes as a function of client size. Previous studies in OTC markets relied on the number of dealers a client trades with as their measure of network size, as trade-level data allowed them to only observe relationships that led to trades. My dataset allows me explore two alternative definitions of a client's network, which contain richer information on the actual trading opportunities that a client faces. First, I can measure the network size as the number of dealers that a client contacts, i.e., requests a quote from; second, I can measure the number of dealers from which the client actually receives a quote. As noted above, there is a substantial difference between network sizes defined on the basis of trades versus quotes, as of the 9,453 dealer-client relationships, only 59% (5,624) end up leading to a trade. Longer time series should lead to a convergence of these two measures (based on

[7]72.5% of all RFQs (by volume) lead to a trade; however, the non-execution of 27.5% of RFQs is driven mainly by banks (Table 1.1). Many of the non-executed RFQs are due to match makers, discussed later. A match maker, upon being contacted by a client, immediately contacts other dealers to unload the trade, and she trades if and only if the client trades with her. Large clients frequently contact match makers, but seldomly trade with them. Therefore, many RFQ appear not to lead to trades

provision of quotes and based on trades); however, relying on the convergence of the number of dealers that a client trades with to the number of dealers providing a quote comes at the cost of measuring network changes at a very slow pace, as one cannot precisely infer when no more quotes were being provided based on trades alone. Measuring the network size via the number of dealers the client contacts may overestimate the number of trading opportunities a client has, as some dealers may not provide a quote to the client. The measure is still informative as it allows to infer how easily and willingly dealers provide quotes to clients. Ultimately, the primary definition I adopt for measuring the size of a client's network is the number of dealers that provide the client with a quote.

Figure 1.1 shows that a client's network size is strongly related to the size (trade volume) of the client. I classify clients as small, medium or large, based on their trade volume, with small clients trading less than 10 mil (EUR), medium clients trading 10 to 100 mil and large clients trading above 100 mil during the sample period. Among small clients, 40% receive quotes from exactly one dealer, whereas only a few (<10%) receive quotes from more than 10 dealers. For medium clients, who represent 30% of the sample, only 20% of clients receive quotes from only one dealer, and roughly 30% of clients receive quotes from more than 10 dealers. For large clients, a large network size is the norm rather than the exception; 60% of large clients receive quotes from more than 10 dealers. Figure 1.2 shows that these distinct patterns are robust to the definition of network size. These results are more formally summarized in Table 1.2, which shows the relationship between client network size and client size.

The fact that smaller clients receive fewer quotes from dealers compared with larger clients, as observed in Figure 1.1, aligns with prior findings regarding the relationship between client size and network size (e.g., Hendershott et al. (2020a)). Notably, my data further enable me to distinguish whether this pattern arises because dealers do not provide quotes to (small) clients (representing a dealer-enforced restriction), or because small clients decide to contact few dealers (representing a (constrained) choice by small clients). In Figure 1.2, I observe that the left and middle columns are very similar, suggesting that there is little difference between the number of dealers that are asked to provide a quote and the number of dealers that provide a quote. In other words, clients receive quotes from almost all dealers they contact, and their smaller network size is an outcome of their choice to contact only a few dealers.

These findings are further supported by Figure 1.3, which shows that, in most cases, dealers indeed provide quotes to clients that request them. Each set of bars shows a different way of measuring the probability of a dealer providing a quote if a client requests a quote. The leftmost column measures the dealer response rate at the client-dealer-currency pair level, with each observation being the probability of a dealer providing a quote to a client in a currency pair. The middle column measures the dealer response rate at the client-dealer level, pooling across the different currency pairs. Finally, in the rightmost column, the dealer response rate is measured at the client level, pooling across all clients and dealers. All columns show high response rates by dealers (above 80% of quotes requested lead to a quote being provided by the dealer). The differences across client sizes are economically small.

The findings in this subsection thus indicate that small clients receive quotes from fewer dealers (i.e., have smaller client networks) compared with large clients, and, importantly, that these differences in network size are driven by small clients choosing to contact fewer dealers compared to large clients, rather than by dealers responding less to small clients.

There are caveats to this finding, as the data do not capture cases in which a small client avoids approaching a particular dealer because of prior knowledge that the dealer will not respond (e.g., because the dealer does not accept small clients). While I cannot rule out the possibility that such behavior shapes client network sizes, in the following section I will provide evidence consistent with the idea that small clients are free to contact many dealers (with the reasonable expectation of a response), yet choose to not do so. It is also worth noting that 47% (503) of the small clients received quotes from the Top 15 dealers, and 64% (320) of those clients received quotes from at least two Top 15 dealers. Similarly, more than 20% of small clients received quotes from at least 6 dealers. These findings suggest that, despite their low trade volume, many small clients can receive quotes from large/many dealers[8].

1.4.2 Transaction Cost by Client Size

The previous section showed that, compared with large clients, small clients choose to have much smaller client networks, i.e., receive quotes from fewer dealers. In this section I explore how client size and the size of a client's network relate to the transaction cost a client realizes. Auction theory predicts that auctions involving more bidders lead to higher revenue for the seller; in my case, this logic implies that a larger client network—i.e., more bidders—should lead to lower transaction costs. In Table 1.3, I study how the transaction cost of a trade is influenced by client size and client network size.

In the equations that follow, i refers to the client initiating the RFQ, j is the dealer the client is trading with, t measures the time, and c indicates the currency pair. Transaction cost (TC) is measured as the difference between the Dealer's Quote, generally the dealer executing the trade, and the midquote in log basis points[9] (bps):

$$\text{TC}_{i,j,c,t} = \left(1 - 2\text{Sell}_{i,c,t}\right) \times \Big(\log(\text{Dealer's Quote}_{i,j,c,t}) - \log(\text{midquote}_{c,t})\Big) \times 10^4 \tag{1.1}$$

In examining how transaction cost relates to client size (*Client Size*) and to network size, I adopt two alternative definitions for the client network: the inverse of (i) the client network for

[8] Di Maggio et al. (2017) show that core dealers (large dealers) have much bargaining power, as they were able to charge much higher prices (to peripheral dealers and clients) during the financial crisis than peripheral dealers. Thus, if indeed dealers curtail access to small clients, I would expect large dealers to be particularly likely to do so. Since many small clients contact not just one but many large dealers it appears not to be the case that large dealers restrict access to smaller clients.

[9] Transaction costs are winsorized at the 1- and 99-percentile. The 99-percentile is around 10 log basis points, and therefore the difference between basis points and log basis points is negligible.

the specific RFQ (i.e., the number of dealers who provide a quote for the RFQ; *1/Dealers in RFQ*); and (ii) the client network throughout the data collection period (i.e., the number of dealers that provide a quote across all RFQs; *1/Client Network Size*). An additional specification controls for the number of quotes in the RFQ, via the Herfindahl–Hirschman Index of the concentration of a client's total trading volume among the dealers (*Client HHI*). The regressions include currency pair and time fixed effects (in the form of *Date* and *Time of Day*). Further, I control for the client type (e.g., corporation, institutional investor, etc.), trade size, and whether the trade direction was revealed before the trade. Thus, I estimate the following panel regression:

$$\begin{aligned} \mathrm{TC}_{i,j,t,c} &= \alpha_t + \alpha_c + \left(\alpha_j\right) + \beta \log\left(\text{Client Size}_i\right) \\ &\quad + \gamma \left(\text{Dealers in RFQ}_{i,j,t,c}\right)^{-1} + \vartheta X_{i,j,t,c} + \epsilon_{i,j,t,c} \end{aligned} \tag{1.2}$$

In Table 1.3, Specification (1) shows that larger clients, measured by larger trading volume, trade at lower transaction costs, with a client 10% larger receiving 0.05 bps(=$0.52 \times \log(1.1)$) lower transaction cost, which is a reduction of 6.1%(= $\frac{0.05}{0.82}$) of the average transaction cost. The controls for trade size show that transaction costs are U-shaped, with the smallest trades having the highest transaction costs, and larger trades having at first decreasing and later increasing transaction costs, which is typical for OTC markets (Edwards et al., 2007).

In Specifications (2) through (4), I add the client's network size to the regression. The network is best approximated by *1/Dealers in RFQ*, as the marginal benefit of adding a dealer is diminishing in the number of dealers. Specification (2) shows that having a small network is particularly costly, with the average trade cost being 1 bps larger if the client receives a quote from one instead of two dealers. Furthermore, the inclusion of network size accounts for 30% of the price improvement that larger clients receive. Specification (3) shows that controlling for the network size, either via the number of dealers providing a quote in the RFQ (Specification (2)) or via the number of dealers that provide a quote to the client at some point in time (Specification (3)), does not change the importance of client size; nor does the change in measurement qualitatively change the relationship between network size and transaction cost.

Specification (4) controls for the network size via *Client HHI*, the Herfindahl-Hirschman Index of of the concentration of a client's total trading volume among the dealers. While the Herfindahl-Hirschman Index does not relate to network size, the dynamics it captures are likely to be similar to those reflected in *1/Client Network Size*. For example, in case where a client trades with all dealers equally, the Herfindahl-Hirschman Index will match the *1/Client Network Size*. Furthermore, higher dispersion of a client's trading across dealers (low *Client HHI*) means that more dealers provide good quotes to the client, which should be important for the trading cost. The inclusion of any network size measure in Specifications (2) to (4) does not change the coefficient on client size in an economically significant way, and the point estimates on the network size measures are very similar across measures, showing that all three measures are an accurate depiction of the network size.

Finally, Specifications (5) through (8) repeat Specifications (1) through (4) but add dealer fixed effects. The coefficient of network size is quantitatively unchanged in Specifications (6) and (7), and the introduction of dealer fixed effects reduces the effect of client size on transaction cost by 50% in Specification (5). Specifications (6) through (8) show that there is some correlation between the dealer fixed effects and the network size, with larger dealers being more prevalent in larger networks; still, this correlation is low, with only 0.05 bps of the 0.25 bps reduction in coefficient on *Client Size* being due to the client's network size. While the inclusion of additional controls, such as client network size or dealer fixed effects, reduces the slope of the relationship between client size and transaction cost, the slope remains negative. Overall, it can be said that 30% of the difference in trading costs between small and large dealers is due to client network size, and a further 40% is due to the identity of the dealer. It can thus be concluded that larger clients receive lower transaction costs through at least two channels: 1) large clients have larger networks; and 2) the networks of larger clients are more likely than the networks of smaller clients to include larger dealers, who provide lower transaction costs.

The results indicate that contacting more dealers is always beneficial for a client, suggesting that any client can achieve lower transaction costs by requesting quotes from more dealers. Yet, the results outlined in the previous subsection indicate that smaller clients choose not to reach out to many dealers, despite the fact that dealers typically provide quotes easily to clients. Together, these findings indicate that reaching out to large numbers of dealers must come at a cost to clients; otherwise, they would do so and thus receive lower transaction costs. Such costs might include unobservable network formation/maintenance costs on the client's side—yet, in what follows, I reveal additional costs that arise when a client contacts and trades with a large number of dealers, and that can explain clients' reluctance to reach out to large numbers of dealers.

Table 1.4 introduces the *Client Size/Client Network Size*$_i$ ratio, the average trade volume between a client and a dealer providing a quote to a client. The inclusion of *Client Size/Client Network Size* reduces the effect of client size, so it is less than one tenth of the original effect once dealer fixed effects are included, and the client now faces a real trade off when selecting the network size[10]. In Specifications (2) and (4), if N is the client's network size, then transaction cost, as a function of network size, is given by:

$$TC_i(N) = -0.5702 \times \frac{\text{Client Size}_i}{N} - 0.8831 \log(N) \tag{1.3}$$

[10]The specification was changed to account for the fact that trade volume per dealer is important for transaction cost. To achieve differential optimal networks under such a specification, the Client Network Size cannot enter via the same functional form for client size and independently as network size. Furthermore, both bilateral trading with a dealer and network size should enter negatively and such that $N = \gamma \times \textit{Client Size}$. These choices demand a functional form, where the first-order condition solves for a maximum when taking the model seriously. As the focus here is on the partial derivatives and not the optimal choice of network for each given client , the given specification was chosen to give ease of interpretation.

Thus, increasing the number of dealers leads to a reduction in trading cost, but only if the trade with the average dealer is sufficiently large. Table 1.4 thus shows that the size of the client is not the only important factor in achieving good prizes; rather, the trade volume with each dealer is also important. This aspect is discussed in depth in the following subsection.

1.4.3 Role of Bilateral Trading Volume in Transaction Costs

To further highlight the importance of bilateral trading between clients and dealers, I introduce three further pieces of evidence that support the idea that transaction costs are influenced by the trade volume per dealer (bilateral trade volume) and not the client size, such that a client that trades more with a dealer receives better prices than a client that trades less, even if the latter client is larger.

First, Table 1.5 compares within an RFQ the quotes that are provided to a client across dealers who respond to the client's RFQ, as a function of the bilateral trade volume between the client and the respective dealer. To that end each specification includes RFQ fixed effects, comparing the quotes provided to a client by dealers at one point in time. Specification (1) shows a clear negative relationship between the trade volume between a client and a dealer and the transaction cost the dealer charges. Specification (2) shows that this finding is robust to dealer fixed effects, so it is not simply a dealer providing good prices to all clients. Furthermore, the point estimate is largely unchanged, showing that the price provided by the dealer is affected by the identity of the dealer to only a very small degree[11].

While Specifications (1) and (2) show that there is a relationship between the bilateral trade volume between a client and a dealer and the quote that a dealer provides, they do not rule out the possibility that the relationship is not causal and may instead be spurious. For example, a dealer may provide the same quotes to all clients, but the dealer's willingness to trade may vary randomly over time. In this case, clients that trade extensively with this dealer might simply be those clients that happened to contact the dealer at times when the dealer was providing good quotes. To rule out this possibility, I therefore focus solely on RFQs that do not lead to trade. In these RFQs, the client does not increase the trading volume with the dealer, and thus if the relationship persists it is not spurious, but causal. The results can be found in Specification (3) and (4) with point estimates that remain highly significant and very much of the same order of magnitude. These findings show that bilateral trade volume is important in determining the price provided to a client. Clients receive better prices from dealers with whom they trade more.

Second, I obtain additional support for the importance of bilateral trading between a client and a dealer by showing not only that clients receive better prices from dealers with whom they trade more compared to dealers with whom they trade less, but also that a dealer provides better prices to clients with whom the dealer trades more, independent of the client's size. Thus,

[11] This analysis is conditional on the client contacting multiple dealers, as the RFQ fixed effect subsumes any single-dealer RFQ.

not only is bilateral trade volume between a client and a dealer an important factor in the price provided by the dealer, but it even outweighs the client size in importance[12].

Table 1.6 shows how the prices provided to two clients by the same dealer for the same quantity and at the same point in time differ. For this analysis I match each quote that a dealer provides in an RFQ to all other quotes that this dealer provided in other RFQs at the same time, for the same currency pair and same quantity. For all such matches, I compare the prices offered to each pair of clients, reporting whether the larger client receives a better price (>0), the same price (=0), a worse price (<0), or one or both parties receive no quote, conditional on both the smaller and the larger client trading with the dealer.

The first row shows that the larger client, on average, receives better prices than the smaller client. However, the second and third rows, which distinguish the clients according to their bilateral trading volume with the specific dealer, show that the size of the client is not what matters. While the larger client is more likely (75%) to be the client trading more with the dealer providing the quote, there are many cases in which the smaller client trades more with the dealer. In these cases, the smaller client receives a better price from the dealer than the large client does. Thus, a client that trades more with a dealer receives a better quote from the dealer in 41-49% of cases, while the client that trades less with the dealer receives the better quote in only 15-18% of cases, independent of which is the larger client. Therefore, in order to receive a good price from a dealer, a client must trade at high volume with the dealer offering the quote.

Third, I assess the economic magnitudes of the differences in spreads based on the bilateral trading between a client and a dealer. For this analysis, I rely on cases where dealers provide both a buy and a sell quote in an RFQ. When both a buy and a sell quote are provided, I can control for the inventory of a dealer, which is not observed, by reporting the spread the dealer charges a client. The spread quoted by a dealer varies much less than the quotes used in the previous analysis, as inventory held by a dealer improves one quote, but worsens the other quote, likely in a symmetric effect. Therefore, in this analysis all provided quotes are used and not just quotes that lead to a trade. The dependence of this spread on client size and bilateral trading between a client and a dealer is reported in Table 1.7. For clients that trade with a dealer, the quoted spread, the difference between ask and bid in log basis points, is reported and should be interpreted as being twice the size of the half-spread reported before. All specifications show that the effect of trade size, when divided by two due to the different spread measurement, are very much the same as reported before. However, a strong difference in client size can be observed. While client size has almost no effect on the quoted spread (Specification (1)), bilateral trading between a client and a dealer strongly negatively predicts the spread quoted by a dealer, in a very similar magnitude as in Table 1.5 ($-0.075 = \frac{-0.15}{2}$ vs. -0.06). Specifications (3) through (4) control for the number of dealers in the RFQ, while

[12]Client size and bilateral trading with a dealer are of course correlated; larger clients trade more with a dealer, on average. However, I show that it is the bilateral trade volume between a client and a dealer and not the size of a client that determines the quote provided by the dealer.

Specification (5) controls for the client, all confirming that the bilateral trading between a client and a dealer is the main determinant of transaction cost and not client size.

1.4.4 Implications for the Formation of Networks

The results presented above show that in order for clients to receive good prices in the FX market, clients need to contact many dealers and trade to a substantial degree with each of these dealers. Importantly, as shown in Tables 1.6 and 1.7, a dealer provides better prices to clients that trade more with that dealer, independent of the client's size (i.e., overall trading volume). Therefore, the quote provided by a dealer i to client j at time t can be formulated as the following function:

$$q_{i,j,t}(X) = m_{j,t} + q_j\left(\text{Bilateral Trade}_{i,j}, X, Y_t\right) \tag{1.4}$$

where $m_{j,t}$ is the midquote of the quotes offered by the dealer, which captures the dealer's willingness to trade. The willingness to trade $m_{j,t}$ differs across dealers owing to various reasons, e.g., different inventory positions that make some dealers more willing to buy and others more willing to sell. The variable $q_j\left(\text{Bilateral Trade}_{i,j}, X, Y_t\right)$ is the half-spread charged by the dealer, based on (i) the characteristics of the trade, denoted X (these characteristics include, e.g., trade size, currency pair, etc.); and (ii) market characteristics, denoted Y_t, which include volatility or spreads in the interdealer market. Finally, the spread is determined by the bilateral trading between a client and a dealer *Bilateral Trade*, with better prices to clients that trade more with the dealer, so that $\partial_1 q_j < 0$, and it is reasonable to assume that the incremental price improvement is decreasing in the amount of bilateral trading between the client and the dealer, so $\partial_1^2 q_j > 0$.

Herein I will not attempt to address the question of why dealers offer smaller spreads to clients that they trade more with, instead taking it as given and deriving implications for the behavior of clients. Bernhardt et al. (2004) and Green et al. (2007) obtained a similar finding, observing that larger trades have lower transaction costs; they argued that these better prices are a result of the fact that larger clients have greater bargaining power. In any case, given that dealers reward more trades with better prices, clients face a trade-off. Specifically, when contacting multiple dealers, a client i will choose the best quote offered. Without loss of generality, assume the client wants to buy:

$$\begin{aligned} q^{\star}_{i,t}(X) &= \min_{j \in N_i(X)} q_{i,t,j} \\ &= \min_{j \in N_i(X)} m_{j,t} + q_j\left(\text{Bilateral Trade}_{i,j}, X, Y_t\right) \end{aligned} \tag{1.5}$$

Increasing the number of dealers will always lead to having dealers more willing to trade (lower $\min_{j \in N_i(X)} m_{j,t}$). However, at the same time, the bilateral trade with each dealer decreases the more dealers a client receives a quote from and trades with, so $q_j\left(\text{Bilateral Trade}_{i,j}, X, Y_t\right)$ increases in the number of dealers. Thus, clients increase the number of dealers as long as the expected reduction in transaction cost in finding a dealer more willing to trade is not outweighed by the increased cost of trading less with each dealer.

The transaction cost reduction of finding a dealer more willing to trade should mainly be constant across clients; in particular, it is constant for clients who have the same network and are considering adding the same dealer to the network. However, the costs of trading less with each dealer are different across clients, such that larger clients, who can continue to trade to a more substantial degree with their existing dealers (compared with smaller clients), face lower increases in spreads offered to them by their dealers. Therefore, larger clients choose to have larger networks than smaller clients, as their larger trading volume allows them to continue to reduce transaction costs even as they trade with more dealers. This hypothesis is confirmed in Figure 1.1 and in Figure 1.2, which show that small clients choose to contact fewer dealers.

Small clients, in turn, trade little with even a single dealer. Increasing the number of dealers may not lead to lower transaction costs for such clients, as the reduction in trading with an existing dealer may outweigh the benefit of trading with a dealer more willing to trade. Some small clients might prefer an "even smaller" network, something that is confirmed in Table A.2 in the Appendix A.2, where I study how client size relates to transaction costs of clients contacting only one dealer. There, I find that the relationship between client size and trading cost is even stronger than in Table 1.3, where all clients are studied. The problem for small clients is that they appear unimportant to a dealer and as if they trade only little with them, which is true.

1.5 Match Makers

Section 1.4 showed that small clients receive worse prices from dealers because they trade little with those dealers. Furthermore, to ensure that they receive the best possible price from a dealer, small clients need to do all their trading with a single dealer, preventing them from trading with the dealer most willing to trade. However, the results suggest that if small clients could pool together, and avoid the appearance of being small by trading under one "identity", they could receive better prices from a dealer. The same "identity" could then also contact more dealers, as its larger size reduces the cost of adding dealers to a network. In this section, I show that such a service exists, and that it is particularly targeted at small clients. This service, which, to my knowledge, has not been documented, is provided by dealers whom I call "match makers", described in detail in what follows.

1.5.1 Definition of a Match Maker

A match maker is a dealer whose sole function is to match clients to other dealers. In all cases in which a client contacts a match maker, the match maker immediately passes on the trade to another dealer through an RPT, thereby, allowing the client to indirectly trade with the other dealer. As the match maker initiates a separate RFQ, the dealer does not learn about the identity of the client; in other words, all clients of the match maker appear under the same (i.e., the match maker's) name. Furthermore, in the process the match maker never takes hold of any inventory, in my case exposure to any currency pair. The match maker charges a fee between the price provided to her by the dealer and the price charged to her client.

Panel A in Table 1.8 shows an example of a typical trade with a match maker. First, *Client 1* contacts match maker *MM 1*, as can be seen in the earliest Request ID being given to the RFQ initiated by *Client 1*. After being contacted by the client, the match maker immediately contacts her dealers (*Dealer A-D*) for the same trade, as observed in the fact that the currency pair, trade direction and trade size match those of the client. The immediacy in initiation can be seen by the Request IDs for the client and the match maker being very close to each other. After receiving quotes from her dealers, the match maker refers the best quote to client for a fee; in Table 1.8 the best quote is quote 1.08979, to which a fee of 0.5 pips is added, so that 1.08984 is offered to the client. Once the client decides to trade, the match maker executes a trade with her dealer (*Dealer C* in Table 1.8), and utilizes her "last-look" option, delaying the trade confirmation until her trade has been confirmed by *Dealer C*. The delay of trade confirmation can be seen through the differences in Tradetime, with the match maker trading with her dealer before trading with her client, a difference of 0.036 seconds. This immediacy and delay of trade confirmation makes the matching of the client to one of the match maker's dealers riskless.

Furthermore, Panel B of Table 1.8 shows an example where a client contacts a match maker (*MM 2*) for a trade, but does not execute the RFQ. Most aspects of the process are the same as in the example provided in Panel A: The client initiates the RFQ first, with the match maker following immediately thereafter; the RFQs are still for the same trade, and the match maker still charges a markup over her best-quoted price (in the current example, the markup is 0.8 pips, the difference between *Dealer H*'s quote and *MM 2*'s quote). The only difference is that the Tradetime, the time the RFQ ends, is now earlier for the client than for the match maker. When there is no trade, it is the client, and not the match maker, who decides when the RFQ should end. Therefore, the match maker only learns about the end of the client's RFQ once the client cancels the RFQ or lets the RFQ expire. Only then can the match maker react to the client's action, canceling her RFQ immediately, 0.015 seconds later. Appendix A.4 confirms that the examples in Table 1.8 with respect to the timing differences are not exceptions, but the norm, as most RPTs that culminate in trades have the match maker trade before her client, and most RPTs that end without trades have the client end the RFQ before the match maker. Furthermore, despite the fact that I identify matching RPTs that take place within a 3-second window, most RPTs that I identify have Tradetime differences of less than 0.2 seconds.

Panel B of Table 1.8 further confirms that trades between clients and match makers are not prearranged. If the trades were prearranged, the match maker would be contacted only when a trade takes place, and thus there would be no unexecuted RFQs between a client and a match maker or RFQs in which the match maker initiates an RFQ and request a quote from her dealers, but does not trade with the client and the dealers. The lack of pre-arrangement contrasts the RPTs most often studied in corporate bond markets, as in the work of Saar et al. (2019), where dealers are contacted by a client and search for another client to trade with. The match maker therefore reacts to the trading needs of her clients, and her activity is purely a reaction to her clients wanting or not wanting to trade, conditional on the quote provided. There are no instances in my data in which a client does not want to trade but the match maker nevertheless executes her side of the RPT[13]; likewise, in all cases in which the client decides to trade with the match maker, the match maker subsequently executes her side of the RPT.

It is important to acknowledge that any dealer can trade via RPT and may have good reason to do so; for example, if a client contacts a dealer for an exotic currency pair, the dealer may not have the overall client demand to find another client, or the expertise to price the currency pair correctly, and may thus decide to utilize an RPT to provide the client with a quote. Hedging and risk-sharing may further motivate a dealer to pass on a significant proportion of her trades to another dealer in a short amount of time. Thus, I define a match maker as a dealer that exclusively utilizes RPTs that cover the complete trading size and (almost) never takes any exposure to a currency pair. Furthermore, a match maker will not trade on her own account and only acts to match clients to dealers (client-dealer intermediation).

1.5.2 Prevalence of Match Makers among Dealers

Identifying match makers will be restricted to match makers that utilize the focal multi-bank platform for all aspects of the match making role, namely, receiving clients' requests and requesting quotes from other dealers. Given that match makers might use other platforms to perform part of the client-dealer intermediation (e.g., they might receive RFQs on the multi-bank platform and unload the trade to another dealer on a separate single-bank platform, or vice versa), my data provide a lower bound on the prevalence of match makers among all dealers.

Appendix A.3 describes how I manually identify, among the dealers in my data, those who provide quotes and request quotes from other dealers on the platform. Specifically, I find that out of 334 dealers, 97 dealers also request RFQs on the platform. Of those 97 dealers, 88

[13]There are 22 out of over 18,000 RPTs in which the match maker appears to execute a trade with her dealer, but the client does not trade. These cases are due to clients contacting the match maker and other dealers for a trade, deciding to execute the trade with a non-match maker dealer, and immediately submitting another identical trade that gets executed with the match maker. The match maker will provide a quote to both RFQs of the client, but she herself will have only one RFQ with her dealers, as her initial RFQ is still on-going and was not canceled before the client's second RFQ began.

dealers execute trades in RFQs, and for each of those 88 dealers, I look at the share of their "total trading" that comprises RPTs in RFQ:

$$\begin{aligned}\text{RPT Volume in RFQ}_i \div \big(&\text{Non-RPT Volume as Dealer in RFQ and Streaming}_i\\ &+\text{Non-RPT Volume as Requester in RFQ and Streaming}_i\\ &+\text{RPT Volume in RFQ}_i\big)\end{aligned} \tag{1.6}$$

Total trading includes, in addition to RPTs in RFQ, the provision of quotes and trading with clients that are not part of an RPT (*Non-RPT Volume as Dealer*) and trading with other dealers that are also not part of an RPT (*Non-RPT Volume as Requester*). Non-RPT trading of a dealer includes not only trading in RFQ, but also via Streaming, to ensure that a dealer that only does RPT does so across all trading mechanisms and for all trades. The distribution of RPT as a share of total trading is depicted in Figure 1.4.

Figure 1.4 shows what fraction of the distribution is bimodal, with most dealers executing almost exclusively RPT ($\geq$ 95%) or very little RPT ($\leq$ 35%); in the middle part there are relatively few dealers. Given this bimodal distribution, I classify the top 2 groups ($\geq$ 85% of trading is RPT) as "match makers"[14]. Thus, I identify 46 match makers out of 97 dealers that also request trades as clients. Assuming conservatively that no other dealer is a match maker, 14%($=\frac{46}{334}$) of the dealers in my data are match makers.

1.5.3 Importance of Match Makers

Table 1.9 shows the summary statistics of match makers. The trade volume shows that match makers execute 99%($=\frac{3,800.13}{3,843.52}$) of their trade volume as dealers via RPT, and that 95% of their trades with other dealers are RPTs. The small difference in volume may be attributable to trades from clients contacting the dealer outside the platform. Compared with all dealers, as presented in Table 1.1, match makers are not the largest (Top 40) dealers, but cover much of the remaining distribution, with 25% looking like the average 201+ dealer, median like a lower 101-200 or higher 201+ dealer, and 75% like a smaller 51-100 dealer or a larger 101-200 dealer. When it comes to the number of currency pairs and clients, it appears as if match makers are slightly smaller than regular dealers, with the means for both variables representing smaller 101-200 dealers rather than larger 101-200 dealers, as is the case for trading volume. In terms of their trade volume as clients, match makers on average represent a larger 191-500 client

[14]Matching by hand reveals that a further 7 dealers should be added to this group. Two of these dealers do not trade at all, but still contact other dealers when they are contacted. One other dealer has a large number of smaller trades (<100,000 EUR), but utilizes RPT for any trade above 100,000 EUR and periodically passes the small trades on in aggregated form. One dealer uses RPT for all transactions, but initiates RFQ in exotic currencies that are not RPT. One dealer executes more than 13,000 RPT, and only 13 RFQs are not RPT. Two dealers execute all client RFQs via RPT, even those that do not lead to a trade, and has minor trading in other currency pairs. The results are robust to the exclusion of any of these dealers

with respect to trading volume, and receive quotes from and trade with as many dealers as an average 191-500 client.

Furthermore, clients seem to trade with at most one match maker. In total, match makers have 214(=46 × 4.65) relationships with clients in which they trade, and these stem from 202 clients, showing that only 6% of clients trade with more than one match maker. That clients only trade with one match maker is a key difference between match makers and regular dealers, as clients generally contact multiple regular dealers. In particular, assuming random network choice, and an average client trading with 3.34 dealers, the average client trading with at least one match maker should trade with multiple match makers in 16%($=1-\frac{\text{Trades with exactly 1 MM}}{\text{Trades with at least 1 MM}}$ $=1-\frac{3.3\times\frac{46}{334}\left(\frac{334-46}{334}\right)^{2.3}}{1-\left(\frac{334-46}{334}\right)^{3.3}}$) of cases. This difference continues to hold when considering the number of match makers whom clients contact with RFQs and from whom they receive quotes: In this case, too, clients generally contact only one match maker.

More than 10% of clients trade with a match maker, and together, match makers execute 6%($\approx\frac{3{,}699.21}{76{,}535.9-9{,}917.69}$) of all client trading volume on the platform[15]. Summing up, I conclude that match makers are popular among clients, and that match makers are highly prevalent on the platform (14% of all dealers).

These findings raise the question of how match makers fit into the larger market network structure—and specifically, whether they are core or peripheral dealers. Figure 1.5 shows in bars the cumulative trading volume of all previous dealers and whether a current dealer is a regular dealer (green bar) or match maker (navy bar). The largest match makers are just within the Top 50 dealers and execute less than 1% of trading volume. Dealers larger than any match maker account for more than 70% of all trade volume. Thus, all match makers are peripheral dealers. These findings show that, in contrast to the common assumption in models of OTC markets (e.g., Wang (2017); Farboodi et al. (2017)), many peripheral dealers perform fundamentally different functions from core dealers, who typically hold inventory from trading with clients and mainly trade with other clients. To my knowledge, this paper is the first to document systematic differences in trading behavior between core and peripheral dealers.

1.5.4 Comparison of Match Makers versus Other Market Actors

In this subsection, I will elaborate on how the match maker's role is distinct from the roles of other, more familiar, market actors. From the client's perspective, trading with a match maker may not appear different from trading with any other dealer, as match makers, like other dealers, provide quotes to clients. However, after trading with a client, the match maker acts very differently from the typical dealer studied in OTC markets, and, specifically, from the large

[15]The match maker RPTs are recorded as two trades, one trade between the client and the match maker and one trade between the match maker and the dealer. Thus at least those trades have to be removed from the trade volume requested by clients, which is 9,917.69. Assuming that all trading by any participant is client trading volume that is not an RPT gives the 6% market share of client trading volume.

dealers that conduct most of the core trading (Li and Schürhoff, 2019): Instead of holding on to inventory and internalizing the trade, as most large dealers do (Schrimpf and Sushko, 2019), match makers immediately pass the trade on, taking hold of no inventory, and thus supply liquidity only indirectly to their clients.

Notably, this intermediary function is not exclusive to match makers: Most dealers in OTC markets serve as both brokers and dealers (broker-dealers), where a dealer is defined as an actor that trades on her own account, while a broker provides best execution to the client, charging a commission to find the best quote, which often leads to trading with another dealer. The dual role of broker-dealer has been criticized by Harris (2015), who argued that clients often do not know whether, in a particular transaction, the broker-dealer is acting as a broker or as a dealer, leading to cases in which the broker-dealer acts as a dealer and provides a worse price than is available at another dealer the broker-dealer has access to. In this framework, the match maker fits the description of a broker, in that she provides trading opportunities with other dealers. However, the match maker has no best execution requirement, and instead can charge clients any fee, in the form of a markup over the price provided to the match maker by her dealers. The match maker thus acts as a dealer to her clients, providing them with prices that she herself chooses.

Another form of intermediary relationship that is prevalent in FX markets is the correspondence banking relationship. In FX markets, a respondent bank that is not able to provide FX trading to its clients on its own is given FX pricing (for a fee) by a correspondent bank through a correspondence banking relationship, providing the correspondent bank with access to its clients (Bank of International Settlement, 2016). Correspondent banks usually have very many relationships with respondent banks, while a given respondent bank typically receives services from only one correspondent bank, including but not limited to FX pricing. While a match maker provides other dealers with access to clients, and receives pricing from those dealers, which is passed on to those clients, match makers are not respondent banks, as they request pricing from many dealers and not just one dealer, as shown below.

Summing up, while match makers are similar both to dealers and to other types of intermediaries we have seen in OTC markets, the manner in which they function distinguishes them from these actors.

1.6 Analysis of Trading with a Match Maker

This section shows which types of clients trade with match makers, whether the match maker provides her clients with additional benefits besides pooling, and whether match makers indeed reduce their clients' transaction costs. This section will focus primarily on the services of match makers to small clients and show that these clients are of major importance to match makers; to a lesser extent, it will also discuss why larger clients may want to trade with a match maker. To guide the analysis, I will develop possible reasons why a particular client depending on his characteristics might seek out the services of a match maker.

1.6.1 Possible Reasons for Trading with a Match Maker

When trading with a match maker, the client does not reveal his identity to the dealers the match maker contacts; instead, those dealers learn only about the match maker. Accordingly, a dealer cannot price-discriminate among different clients of a match maker, and instead needs to treat all match maker clients equally. In addition to avoiding price discrimination, a match maker's clients benefit from the fact that the match maker has multiple clients, thereby leading the clients to appear larger when sharing the identity of the match maker. This benefit is expected to be particularly salient for small clients, such that smaller clients are expected to receive substantially better prices when trading with a match maker than they would receive when trading directly with dealers. Accordingly:

Possible Reason 1.6.1. *A match maker's clients are predominantly smaller (as opposed to larger) clients in the FX market. Furthermore, most of a match maker's clients trade exclusively with that match maker.*

The pooling provided by the match maker to her clients not only makes the clients appear larger but also provides the clients with anonymity. Kyle (1985) shows that informed traders need anonymity to profit from their better knowledge about the value of the currency pair. Accordingly, it is possible that a match maker's clients are informed traders that want to benefit from this anonymity:

Possible Reason 1.6.2. *A match maker's clients are informed traders that utilize the anonymity provided by the match maker to hide the fact that their trades are informed.*

When interacting with other dealers, the match maker appears just like any other client in the FX market. Accordingly, the match maker's network size should be similar to that of any similar-sized client. In turn, the match maker's network provides the match maker's clients with (indirect) access to a different set of dealers than they might otherwise contact. For small clients, who typically have small networks—including clients trading exclusively with the match maker—the (indirect) network provided by the match maker is likely to be substantially larger than the (direct) network that these clients would otherwise be expected to maintain. For larger clients, which typically trade with multiple dealers, trading with a match maker may still be beneficial if the match maker contacts many dealers that the client does not contact, allowing the client to expand his network without needing to form many relationships.

Possible Reason 1.6.3. *Match makers do not differ from other similar-sized clients in their network size. For exclusive clients of a match maker, the indirect network attained through the match maker is expected to be larger than the typical direct network of a similar-sized client who does not use a match maker. Furthermore, for all clients of a match maker, the indirect network is significantly larger than* *these clients' direct networks, and significantly larger than the direct networks that similar-sized clients would be expected to maintain.*

The possible benefits that a match maker provides her clients, outlined above, come at a cost to the client in the form of the fee that the match maker charges on top of the best price she is offered by her dealers. Given the wide popularity of match makers and their relatively large market share, I expect that the match maker provides the client with lower transaction costs than they would otherwise attain through regular dealers, such that the fee charged by the match maker does not outweigh the benefits the match maker provides.

Possible Reason 1.6.4. *The match maker provides her clients with lower trading costs than the clients can achieve on their own.*

The remainder of this section will be used to test the propositions outlined above, and, in particular, to determine whether small and other clients indeed benefit from using the match maker, as compared with other dealers.

1.6.2 Testing of Possible Reasons

1.6.2.1 Small Clients Use Match Makers

Possible Reason 1.6.1 predicted that small (as opposed to large) clients would be particularly likely to use match makers, and that these clients would use a match maker exclusively. Before showing which clients use match makers, I first show that exclusive clients are of particular importance to match makers. The importance of exclusive clients is shown in Figure 1.6, where the distribution of the share of client trading via a single client-dealer relationship is depicted. Most client-dealer relationships cover anywhere from 0 to 35% of a client's trading volume, with less than 15% of trade volume being executed in client-dealer relationships executing more than 95% of the client's trading volume. In other words, less than 15% of all trade volume comes from clients that trade exclusively with a dealer. This distribution is vastly different for match makers. Match makers execute 65% of their trading volume with clients that trade exclusively with them. The remainder of the volume is mainly traded with clients that trade less than 15% of their trade volume with match makers.

Given these findings, I devote particular focus to clients who trade exclusive with a match maker (exclusive clients). These exclusive clients not only make up the majority of the trading with match makers, but also represent the majority of the clients trading with the match maker (108 out of 202). Notably, almost all of these clients—all but one—not only trade exclusively with the match maker but also contact her exclusively.

Having shown that the majority of a match maker's clients contact the match maker exclusively, and that these exclusive clients account for most of the match maker's trade volume, I now test the second part of Possible Reason 1.6.1, namely, that these exclusive clients are small.

Table 1.10 shows how client size is related to the likelihood of trading exclusively with a match maker, i.e., being an exclusive client. Specification (1) shows that larger clients are significantly less likely than smaller clients to trade exclusively with a match maker. Specification (2) confirms these results by grouping clients by size and showing that the larger the client the lower the probability of being an exclusive client, with the point estimate decreasing monotonically the larger the client is. The specification also shows that there are only minor differences for clients outside the Top 200 clients in terms of the likelihood of being an exclusive client (only 1 percentage point difference), while within the Top 200 the probability of being an exclusive client is estimated to be a full 5 percentage points lower than for clients outside the Top 500. As a comparison, 6.5%($=\frac{108}{1,660}$) of all clients are exclusive clients of any dealer. Specifications (3) and (4) apply logistic regressions to capture the bounded nature of the dependent variable and confirm the results and inferences obtained from Specifications (1) and (2).

Summing up, Figure 1.6 and Table 1.10 show that small clients are the main clients of a match maker. Moreover, whereas the average dealer trades less than 15% of its volume with exclusive clients, match makers attribute 65% of their trading volume to exclusive clients. These exclusive clients are particularly likely to be small, as observed in the fact that exclusive clients are much more prevalent outside the Top 200 clients than within them. Together, these findings confirm Possible Reason 1.6.1.

1.6.2.2 Informed Traders

As discussed above, the anonymity provided by match makers may be particularly appealing to informed traders, who benefit by hiding their trading from market makers (Kyle, 1985) and dealers. However, I propose that there are three main reasons to believe that, contrary to Possible Reason 1.6.2, a match maker's clients are not, in fact, informed.

First, though an informed client may benefit from anonymity, this anonymity comes at the cost of the match maker contacting many different dealers. An informed client would like to limit the spread of his private information; thus, when an informed client requests quotes through an RFQ, the client would like to contact one or few dealers, to limit the number of dealers that learn about the trading. Match makers, however, contact many dealers thus greatly increasing the probability of the information staying no longer private.

Second, if a match maker's clients are indeed informed and not worried about many dealers observing the trading, and trade with the match maker to achieve anonymity, then those informed clients might be expected to contact and trade with many match makers, to access additional dealers from whom their identity could be hidden. However, most clients who trade with match makers—not only small clients but also larger clients—contact only one match maker, and not multiple match makers. Specifically, as shown in Figure 1.6, the 202 clients who trade with match makers form only 214 relationships with match makers that lead to trades. Notably, those clients who do contact multiple match makers also contact additional (non-match-maker) dealers, showing that they do not need to hide their identity.

Finally, a match maker's clients tend to be small, whereas informed clients in financial markets are typically seen as large clients and clients that are consistently informed (Hendershott, Livdan and Schürhoff, 2015; Di Maggio et al., 2019a).

Taken together, the arguments above suggest that a client's wish to hide his identity as a result of being informed is unlikely to be a significant factor driving usage of a match maker's services. Rather, the benefits of anonymity in increasing client's size and network are likely to be more important.

1.6.2.3 Direct and Indirect Network Size

To explore Possible Reason 1.6.3, I first investigate whether and how match makers differ from other clients in OTC markets in terms of their network sizes. I then study the difference in network size both within a match maker's clients—i.e., the difference between a client's direct network and his indirect network—and across clients, comparing network sizes between clients that use match makers and those that do not.

Table 1.11 measures how match makers and their clients differ from regular clients in network size. Specification (1) repeats the previous results as a reference for the other reported results. Specification (1)'s dependent variable is the number of dealers that provide a quote to a client, the (direct) network size. Specification (2) introduces a dummy variable for the match maker and shows that match makers do not differ from other clients of the same size. The specification estimates that match makers receive quotes from an insignificant 0.16 more dealers than do non-match-maker clients of similar size. The similarity between match makers and other types of clients suggests that match makers do not experience more or less difficulty than other clients do in forming or maintaining networks; nor do dealers seem to provide match makers with better prices than they provide similar-sized non-match-maker clients (as, in this case, match makers would have larger networks)[16]. Furthermore, the coefficient on client size is (completely) unchanged.

Specification (3) introduces an additional dummy for clients that receive a quote from a match maker. Again, these clients do not differ from other clients in their network sizes, and the relationship between client size and a client's (direct) network size remains the same.

As previously discussed, the match maker provides her clients with access to an indirect network of dealers, i.e., the dealers whom the match maker contacts via RPT. Specifications (4) and (5) measure the indirect network size of a client. A client's indirect network size is calculated as the number of dealers the client contacts directly that are not match makers, plus the number of dealers that match makers, whom the client contacts, contact via RPT for the client, and whom the client does not contact directly. This definition implies that for clients who do not contact a match maker, the direct and indirect network sizes are the same.

[16]If dealers provide match makers with better prices (at lower quantities) then match makers would experience less of an negative effect through a reduction in trading with existing dealers when increasing their networks, in which case I would expect match makers to choose larger networks compared with other similar-sized clients.

$$\text{Indirect Network Size}_i = \text{(Direct) Network Size}_i - \#\text{Match Maker}_i + \left| \bigcup_{k \in \text{Match Maker}_i} \text{Dealers in RPT of k from i not providing quotes to i} \right| \tag{1.7}$$

While it seems reasonable to expect that a client's indirect network should be larger than his direct network, there is a chance that the indirect network is actually smaller, if the client contacts a match maker who only receives quotes from dealers that the client already receives quotes from directly, so that the match maker does not provide access to any further dealer and instead is a dealer that provides no additional liquidity to the client.

Specification (4) shows that, for clients who receive quotes from match makers, the indirect network is 6.4 dealers larger than the direct network, showing that, in general, match makers significantly expand their clients' networks. This increase in network size holds across all clients that receive quotes from match makers, and may indicate why not all clients trade exclusively with a match maker. In other words, I observe that even large clients with the capacity to maintain their own networks can benefit from a significant expansion in network size—specifically, an increase of 6 dealers—by enlisting the services of a match maker.

Specification (5) investigates the increase in indirect network size for exclusive clients of match makers. An exclusive client contacts only the match maker, and thus has a much smaller (direct) network compared with similar-sized clients who do not use match makers[17]. However, the indirect network provided through the match maker contains 1.5 more dealers compared with the (direct) network of a similar-sized client who does not use a match maker. Thus, the match maker not only provides her exclusive clients with a network that is as large as the client would have chosen on its own without the match maker, but provides her exclusive clients with a network that is larger than the client would have chosen on his own and does so without the client having to form more than the minimal number of trading relationships.

Figure 1.7 provides additional details regarding the manner in which a match maker increases the network sizes of her clients, and further reinforces the consistency across all client sizes. The leftmost column in the figure (*All*) shows that the average RFQ involving a match maker involves 11 dealers, including the match maker, contacted directly by the client. When the match maker is contacted, however, she contacts 9 dealers on average—of which a small number, 2, overlap with the dealers that the client contacts. Thus, the match maker provides access to 7 additional dealers, increasing the client's network of regular dealers to 17.

Column *Trade MM* focuses on RPTs where the match maker executes and passes the trade on. As previously shown in Figure 1.6, most of these trades occur with clients that contact only

[17]Exclusive clients (of match makers) have by construction the smallest possible network. Unreported results show that compared to other similar sized clients, an exclusive client has both a statistically and economically significantly smaller direct network.

the match maker; therefore, in this case, the number of dealers the client contacts directly is smaller, and the overlap between the client's direct network and the dealer's network is also reduced to only 0.5 dealers. Overall, the match maker increases the client's direct network of regular dealers from 4 (asked directly) to an indirect network of 12. The next column (*Exclusive MM*), which focuses clients who trade exclusively with a match maker, shows that the indirect network size achieved via the match maker is similar for these clients.

Columns *Small (<10), Medium (10-100)* and *Large (>100)* return to the sample of column *All* and show how the number of dealers providing quotes differs by client size at the RPT level. For the number of dealers the client contacts directly I find the same pattern previously observed in Figure 1.2, with small clients receiving fewer quotes from dealers than medium or larger clients. A similar difference is also seen when including the number of dealers that provide quotes through the match maker. Importantly, however, the increase in the number of dealers that provide quotes is very much the same across the different client sizes. Small clients receive quotes directly from 1.5 dealers that are not match maker, but around 10 indirectly, so 8.5 more dealers, while medium and large clients increase their numbers of dealers by 8 and 6, respectively. The difference in the number of dealers is driven by the fact that, for a larger client, there is greater overlap between the client's direct network and the match maker's network. Specifically, for small clients the overlap is 0.25; for medium-sized clients the overlap is 1; and for large clients it is slightly above 2.

These findings suggest that a match maker does not discriminate among her clients in terms of the number of dealers she contacts, and any differences in the network expansion she provides are attributable to overlap between clients' existing networks and the match maker's network. Notably, despite such overlap, the match maker significantly expands the networks of even large clients (from 12 dealers, including the match maker, to 17, a 40% increase), suggesting that clients of all sizes can benefit from the services of a match maker. Together, the analyses outlined above confirm Possible Reason 1.6.3.

The results also relate to the literature on endogenous network formation, highlighting that some dealers solely serve as hubs to provide clients indirect trading opportunities with other dealers. Such dealers create a new type of relationship, in which match makers serve as a gateway to many other dealers without providing any additional liquidity, instead sourcing liquidity from other dealers. This service is appealing to clients of all sizes, as it enables even large clients to achieve large increases in network size. These findings provide new perspectives on the study of network formation and centrality calculation. Specifically, they suggest not only that there are additional, generally unexplored relationships that are valuable to all clients, but also that the calculation of network centrality measures may need to be adjusted. The network centrality measures used in the network formation literature (degree centrality, eigenvector centrality betweenness and closeness, see Li and Schürhoff (2019) for an overview) all rely on direct neighbors, and none of them are additive[18], so that indirect relationships such as those

[18]Additive in the sense that the centrality x_i of dealer/client i is $x_i = \lambda \sum_{k \in M(i)} x_k$, where $M(i)$ are the neighbors

involving match makers are not captured, potentially leading certain actors' centrality and network sizes to be underreported.

1.6.2.4 Transaction Cost Improvement

In this section I explore Possible Reason 1.6.4; that is, I investigate whether the benefits provided by a match maker, discussed in the previous sections, outweigh the fee charged by the match maker, so that the match maker indeed provides her clients with lower transaction costs than would a similar-sized dealer who is not a match maker.

In Section 1.4 it was shown not only that smaller clients contact fewer dealers but also that the inclusion of dealer fixed effects reduces the relationship between client size and transaction cost, meaning that larger clients contact dealers that provide better prices. Therefore, larger dealers seem less attainable to smaller clients, suggesting that a client generally faces the choice of contacting a match maker of a certain size or a non-match-maker dealer of the same size. Further, clients may face additional costs to maintain their networks that are unobservable to me, and which are not reflected in the transaction costs associated with network choice. The size of the price difference between a match maker and a similar dealer may therefore serve as an indicator about the network formation and maintenance costs the client faces. To limit the amount of unobservable information, I assume that similar-sized dealers are similarly costly to connect to or are similarly restrictive, and, thus, I compare match makers to similar-sized dealers, controlling for client characteristics, such as client size and client network size.

Table 1.12 shows how the prices provided by a match maker differ from the prices that a similar sized dealer provides a client. The majority of trades a match maker executes are RPT, while a small quantity of trades, accounting for 1% of all trade volume with clients by match makers, are not passed on via RPT. These trades thus do not receive the same benefits as do trades with match makers that are RPT.

Specification (1) shows that, without controlling for client characteristics, the match maker provides slightly better prices than a similar-sized dealer when she executes RPTs, but charges 1.9 bps more in cases when she does not pass the trade on. The average spread (weighted by trade size) for an RFQ with only one provided quote is 2.03 bps; therefore the 1.9 bps addition almost doubles the transaction cost that the client could attain from a similar dealer charging the average spread. This strong increase in transaction cost shows that match makers themselves are not suited for holding inventory, and thus, unsurprisingly, these trades are only a small part of their trading volume. The small volume of these non-RPT trades, however, may suggest that their high transaction costs are ultimately negligible; indeed, non-RPT trades with a match maker have a median trade size of only 14,948 EUR (mean 54,106 EUR), so that trading with a match maker outside of RPT translates into increased cost of only 2.80 EUR (10.12 EUR).

Specifications (2) to (4) now control for the characteristics of the client. The resulting magnitudes are very much the same for each measure of network size control at around -0.4

of i.

bps. Furthermore, the network size controls are all designed to be between 0 and 1, and their point estimates are quantitatively similar. In these specifications, I observe that the match maker still provides worse pricing when not using RPT, but now provides significantly better pricing when executing the trade via RPT. The negative coefficient of -0.39 bps implies that the same trade with a non-match maker would have cost 20% more (mean spread of match maker's RPT trades is 1.52 bps, = $\frac{0.39}{1.52+0.39}$). The increase in the effect size, relative to that obtained in Specification (1), is not surprising, as a match maker's clients tend to be small and also exclusive (to a much greater extent than the clients of other dealers). As larger clients receive better prices on average, and they trade only a fraction with each dealer, a dealer that is similar in size to a match maker, even if it has the same number of clients, has larger clients. It is therefore more surprising that the match maker provides pricing that is on par with that dealer, rather than providing worse prices to her small clients. Given the results of Specification (1) and Tables 1.3 and 1.10 and Figure 1.6, it is therefore not surprising that the match maker provides her clients with much better pricing (20.4%).

Specifications (5) through (7) confirm that these results hold when focusing on the match maker's exclusive clients, which are responsible for most of the match maker's trade volume (Figure 1.6) and tend to be small (Table 1.10). Specifically, these clients obtain transaction costs that are -0.39 bps lower than they would be expected to obtain with a similar-sized dealer. Indeed, this is not surprising, given that, as discussed in detail in previous sections, small clients are particularly well suited to benefit from the advantages a match maker provides, namely, anonymity (and specifically, the capacity to pool together and appear as one larger client) and network expansion.

Overall, these results confirm that match makers help clients—and especially small clients—reduce their transaction costs. That is, the benefits provided by the match maker greatly outweigh the fee she charges, so that the client receives a 20% lower transaction cost when trading with the match maker than he would obtain from a similar-sized (non-match-maker) dealer.

1.7 Conclusion

This study sheds new light on the well-established phenomenon of price discrimination in OTC markets, in which dealers provide better prices to clients that trade more with them, putting smaller clients—who make up the majority of clients in such markets—at a disadvantage. Specifically, using quote-level data from a leading multi-bank platform in the FX market, in which both clients and dealers were identified, I examined how transaction costs are determined and explored means by which small clients might reduce their transaction costs.

First, I characterized the network formation behavior of clients of different sizes, i.e., the number of dealers to which clients submit RFQs. I observed that, though a larger number of dealers is expected to produce lower transaction costs (ceteris paribus), smaller clients choose to contact fewer dealers compared with larger clients. Notably, I ruled out the possibility that this behavior is a result of dealers' lower propensity to respond to RFQs from smaller clients, showing instead that dealers typically respond to all RFQs they receive. Thus, I attributed the limited network sizes of smaller clients to the finding that the price a dealer offers to a client is based on the dealer's trade volume with that specific client—where better prices are given to clients that trade more with the dealer, independent of the clients' overall trade volume. Since a small client, by definition, has low trading volume even if he trades with only a single dealer, he is expected to maintain only a small network comprising very few dealers.

The importance of bilateral trade volume between a client and a dealer, suggests that small clients could reduce their transaction costs if they pooled their trading together and appeared as one larger client. I revealed a class of dealers—previously undocumented, to my knowledge—that provide such a pooling service. Specifically, these so-called "match makers" exclusively function to pool together clients' trades and execute them under their own names by contacting other dealers via RPT, thereby allowing their clients to appear as one. Match makers, which operate on the periphery of the OTC network, are common, with 14% of dealers being match makers and executing 6% of all volume. Match makers are also popular with clients, as more than 10% of clients trade with a match maker.

I subsequently characterized the clients that a match maker serves and the nature of the service she provides. Specifically, I showed that most of a match maker's clients are small, and that 65% of a match maker's trade volume comes from exclusive clients (i.e., clients who trade with her exclusively). For regular dealers, in contrast, only 15% of trade volume, on average, is attributable to exclusive clients. I further found that, in passing on clients' RFQs to her network of dealers, the match maker substantially increases the number of dealers to which her clients have (indirect) access. This expansion of clients' indirect networks is beneficial not only to small clients but also to large clients, with large clients realizing indirect networks that are 40% larger than their direct networks. For exclusive clients, the match maker provides access to 1.5 more dealers than these clients would have been expected to contact in the absence of a match making service. My analyses confirm that these benefits outweigh the fee charged by the match maker, so that match makers provide their clients with (20%) better prices than a similar-sized dealer would. Together, these results suggest that match makers assist small clients in overcoming the price disadvantages they experience, and that these dealers also benefit large clients.

My finding that small clients can, and do, reduce their trading costs by pooling together suggests that solutions that allow small clients to appear like large clients warrant further attention. Indeed, it would be of interest to determine whether additional solutions, similar to the match maker, have emerged in practice. More broadly, it would be appropriate for the providers of such solutions to be transparent about the nature of their services, and to reveal the fees they charge for them.

1.8 Chapter 1: Figures

Figure 1.1: Client Network Size (by Client Size): The figure shows the distribution of client network size for different sized clients. The measure of network size is the number of dealers that provide a quote to a client. Clients are grouped by their trading volume, with dark green clients trading less than 10 mil EUR a month, navy clients trading between 10 mil and 100 mil EUR a month, and dark red clients trading more than 100 mil EUR a month. Each bar *a-b* represents the fraction of all clients in a group (equal weighted) who receive quotes from a number of dealers within the range a-b.

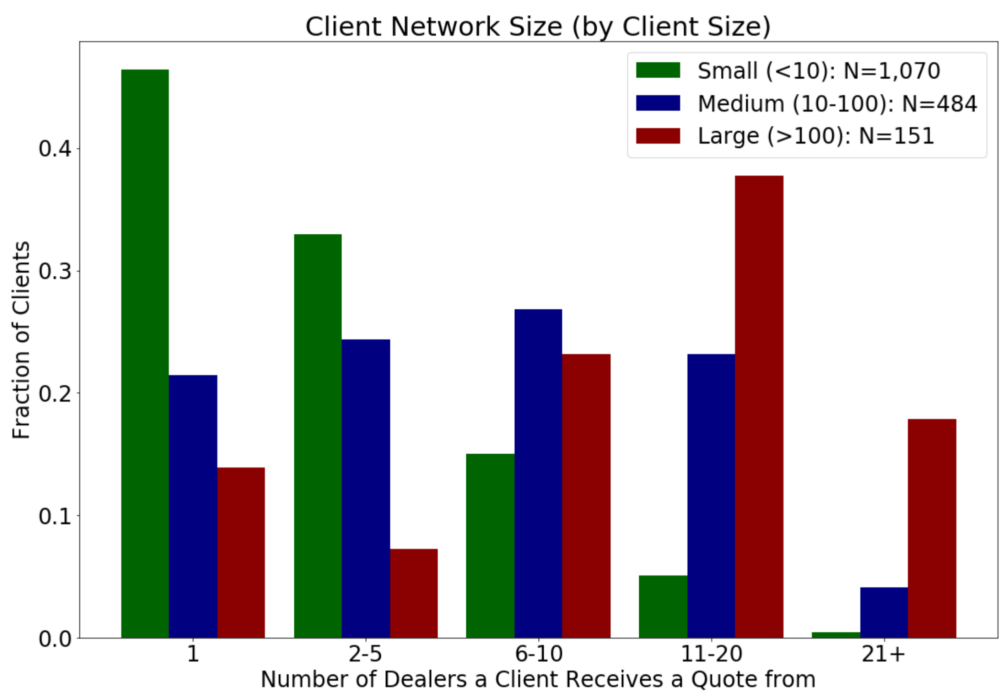


Figure 1.2: Client Network Size across Measurements: The figure shows the average client network size for different sized clients across different measurements of a client's network size. The measure of network size is the number of dealers that a client contacts in an RFQ (left bars, *Dealers (Quote requested)*), the number of dealers that provide a quote to a client (middle bars, *Dealers (Quote received)*) or the number of dealers a client trades with (right bars, *Dealer (traded)*). Clients are grouped by their trading volume, with dark green clients trading less than 10 mil EUR a month (small clients), navy clients trading between 10 mil and 100 mil EUR a month (medium clients) and dark red clients trading more than 100 mil EUR a month (large clients). For each bar the (equal weighted) mean is reported.

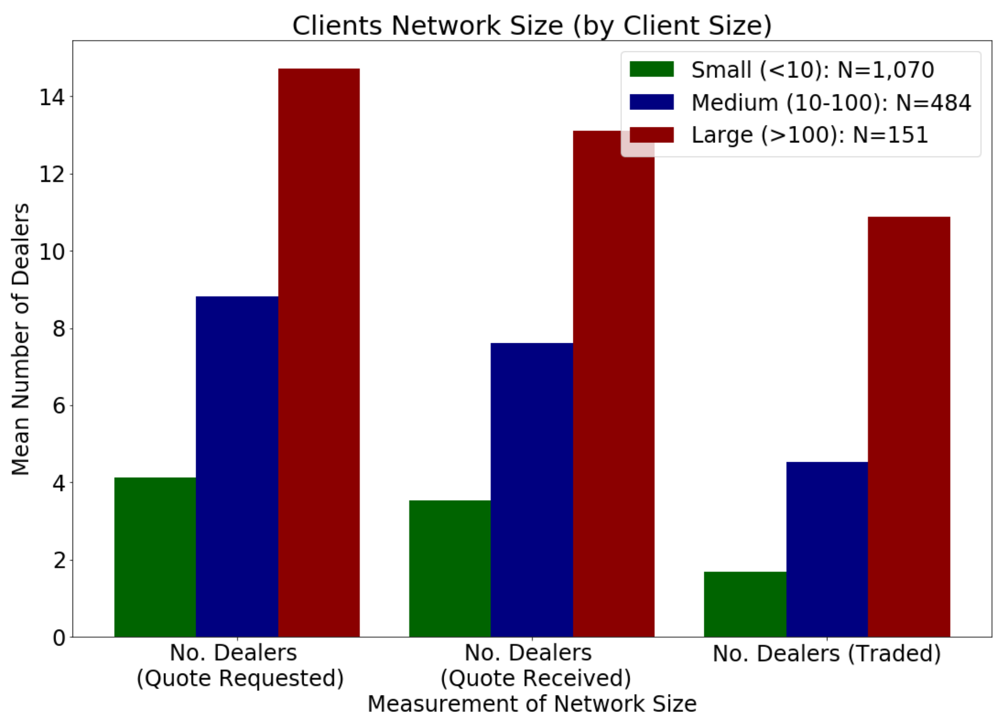


Figure 1.3: Dealer Response Rate (by Client Size): The figure shows the mean dealer response rate for different sized clients. The dealer response rate is measured as the share of dealers a client requests a quote from that provide a quote to the dealer (at some point in time). Clients are grouped by their trading volume, with dark green clients trading less than 10 mil EUR a month, navy clients trading between 10 mil and 100 mil EUR a month and dark red clients trading more than 100 mil EUR a month. For each group the (equal weighted) mean of the dealer response rate is reported. Each group of graphs reports the dealer response rate at a different level. The left bars (*Client-Dealer Currency Pair*) report the dealer response rate at the client-dealer-currency pair level, where each observation reports the dealer response rate of a dealer to a client in a currency pair. The middle bars (*Client-Dealer*) report the dealer response rate at the client level, pooling across currency pairs (weighting each currency by the number of requested quotes within a client), while the right bars (*Client*) report the dealer response rate across all dealers at the client level. Each bar controls for the dealer never providing a quote in the currency pair and the client type.

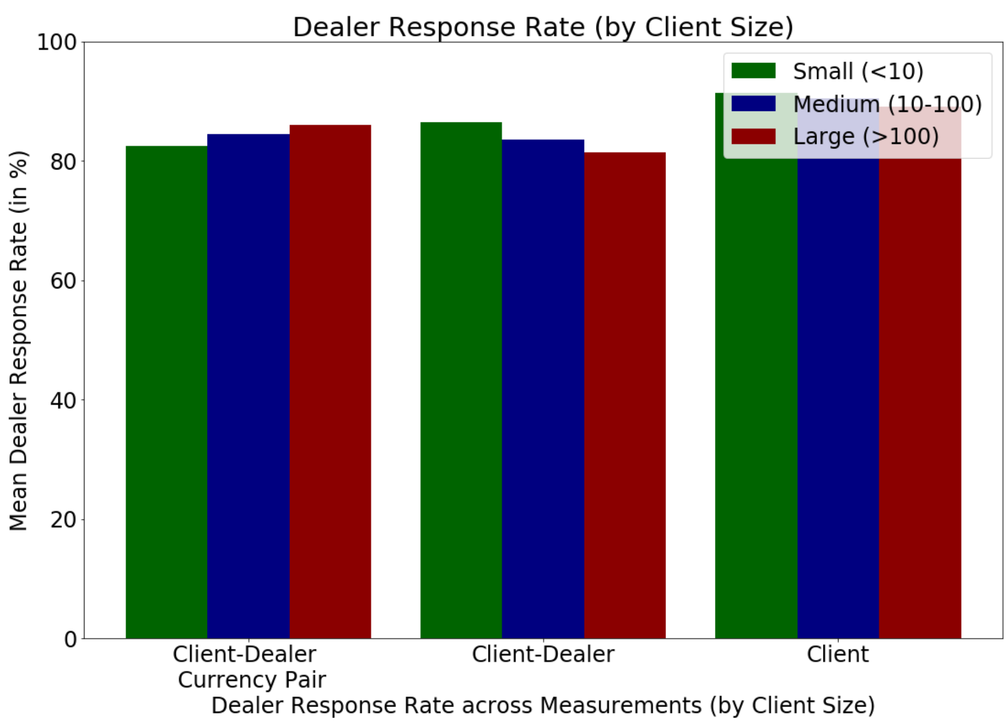


Figure 1.4: Distribution of Dealer's Riskless Principal Transaction Volume as a Share of Dealer's Total Trade Volume: The figure shows the share of a dealer's total trade volume that are riskless principal transactions (RPT). The dealer's total trade volume is defined as the sum of a dealer's trade volume arising from providing quotes to clients (*Trade Volume as Dealer*) and the dealer's trade volume arising from the dealer requesting quotes from other dealers (*Trade Volume as Requester*), adjusted for the double counting of RPT. So the dealer's total trade volume is $Total\ Trade\ Volume_i{=}RPT\ Volume_i{+}Trade\ Volume\ as\ Dealer_i{+}Trade\ Volume\ as\ Client_i$. A dealer's RPT volume as a share of her total trade volume is $\frac{RPT\ Volume_i}{Total\ Trade\ Volume_i}$.

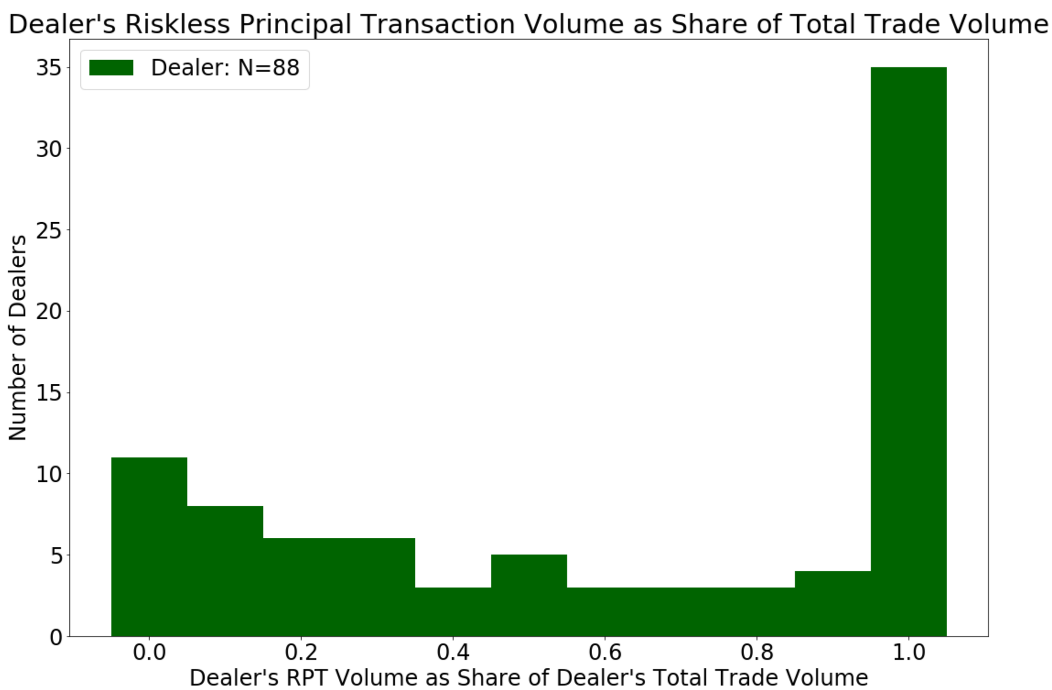


Figure 1.5: (Regular) Dealers' and Match Makers' Market Shares by Rank: The figure shows the rank and the corresponding market shares of (regular) dealers and match makers. Dealers are ranked by size with the largest dealers being left (1) and the smallest dealers being right (334). If the current dealer is regular the dealer is green, while match makers are navy. The left axis provides reference to the cumulative market share of the N largest dealers shown in the green and blue bars. The right axis provides reference to the Nth largest dealer market share, shown in the black line.

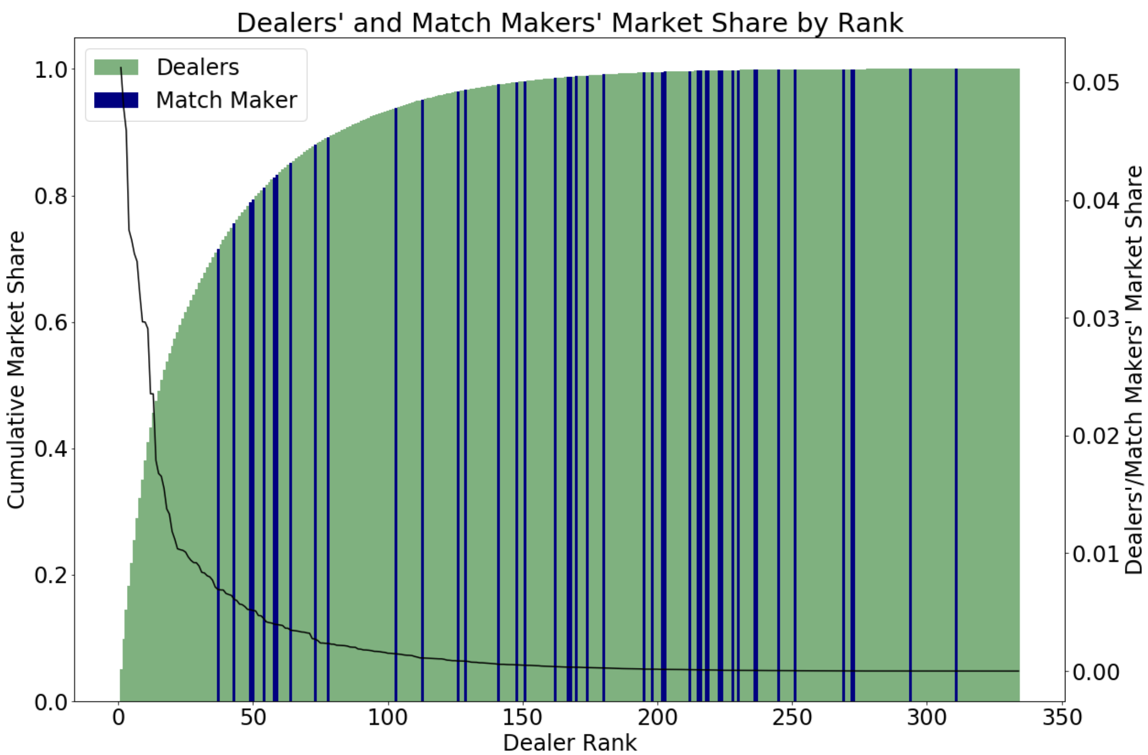


Figure 1.6: Distribution of Concentration in a Client's Trading with Dealers: The figure shows the distribution of the trade volume in dealer-client relationships. The x-axis shows the share of client trading that is executed in a single dealer-client relationship, while the y-axis shows the fraction of all trading volume that is executed in client-dealer relationships that execute a certain share of client-dealer trading. The distribution across client-match maker relationships is shown in navy, while the distribution across all client-dealer relationships (including client-match maker relationships) is shown in dark green.

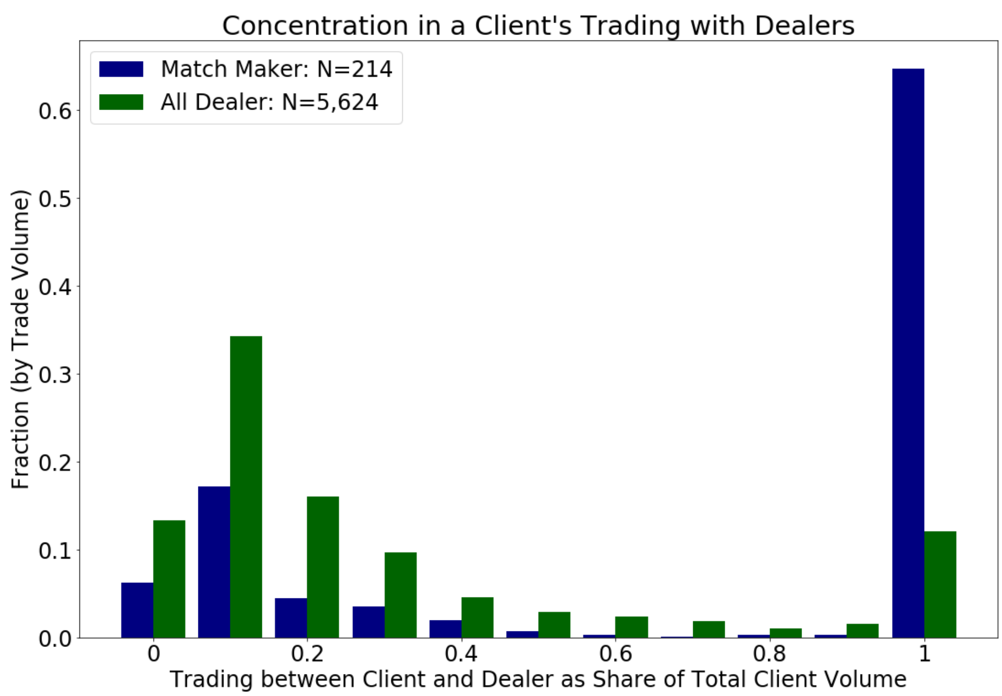
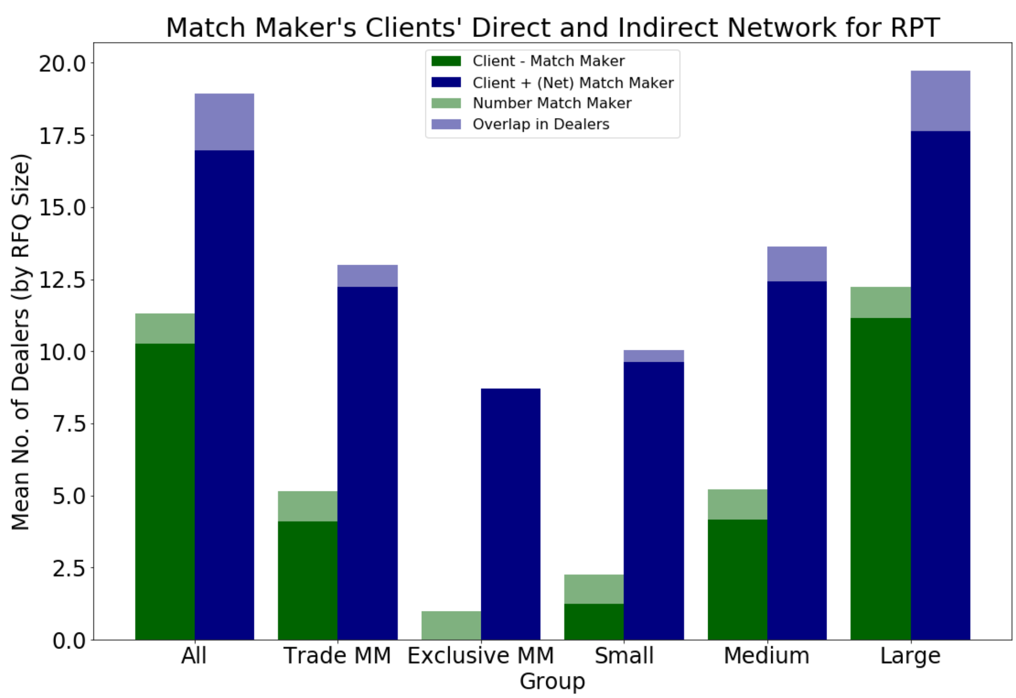


Figure 1.7: Direct and Indirect Network Size when Receiving Quotes from a Match Maker: The figure shows client's direct and indirect numbers of dealers in riskless principal transactions (RPT) involving match makers. The number of non-match-makers a client receives quotes from is depicted in dark green, while the number of match makers is depicted in light green. Together, the light- and dark green show the number of dealers a client contacts directly. Navy represents the number of unique dealers providing a quote to a client either directly or through the match maker. Note that the match maker does not provide a quote, as the match maker (*MM*) simply passes on one of the other dealers' quotes. Lighter blue represents the number of dealers that the match maker and the client contact. Together, navy and lighter blue represent the total number of quotes provided. The y-axis reports the mean weighted by the Trade Size requested in an RFQ (*RFQ Size*), while the x-axis reports the mean of each measure across different groups. The sample is constructed across all RPTs that involve a match maker independent of whether the RPT leads to a trade or not. Different bars include different subsamples of this sample. *All* reports the means for the full sample of RPT, *Trade MM* conditions on RPT, where the client trades with a match maker, and *Exclusive MM* conditions on RPT, where the client only contacts a match maker, but independent of whether the client trades with the match maker. The final 3 columns (*Small*, *Medium*, *Large*) include all RPT where the monthly trading volume by the client is below 10mil EUR (*Small*), between 10 and 100mil EUR (*Medium*) and above 100mil EUR (*Large*) respective, again independent of whether the client trades with a match maker.



1.9 Chapter 1: Tables

Table 1.1: Summary Statistics: Panel A and B: This table reports descriptive statistics for clients in the sample in October 2019. A client is any participant that requests a quote via RFQ. Each column reports the average across all clients in the category, with Panel A focusing on the type of client, while Panel B focuses on client size.*RFQ Volume* is the volume that a client requests via RFQs and receives a quote for. *Trade Volume* is the volume of RFQs that result in a trade. *No. RFQ* counts the number of RFQs a client initiates and receives a quote for.*No. Trades* is the number of RFQs that result in a trade. *No. Dealers (Quote requested)* are the number of dealers a client asks for a quote in an RFQ, i.e., requests a quote from in an RFQ. *No. Dealers (Quote Received)* is the number of dealers that reply at least once to a client in an RFQ, while *No. Dealers (Traded)* is the number of dealers a client executes a trade with. *No. CCY Pairs* is the number of currency pairs for which a client initiates an RFQ and receives a quote (EURUSD and USDEUR are treated as one currency pair). *No. CCY Pairs (Traded)* is the number of currency pairs that a client trades. *Market share* measures the share of total trade volume completed by the group, while Panel B also adds *Corporation, Bank, Institutional Investor*, which report the share of all clients that are a corporation, bank or institutional investor. **Panel C:** This table reports descriptive statistics for dealers in the sample in October 2019. A dealer is any participant who provides a quote in an RFQ. Each column reports the average across all dealers in the category. *(Quote Requested)* reports the variable conditional on a client requesting a quote from a dealer in an RFQ. *(Quote Provided)* reports the variable conditional on the dealer providing a quote to a client in an RFQ. *(Traded)* reports the variable conditional on the dealer executing a client's RFQ. *Volume* sums the volume requested from a dealer by clients in RFQs where the dealer provides a quote (*Volume RFQ (Quote Provided)*), or the dealer trades with the client (*Trade Volume*).*No. RFQ (Quote Provided)* counts the number of RFQs where a dealer provides a quote to a client, while *No. Trades* counts the number of trades a dealer has with clients. *No. Client* counts the number of clients that request a quote from the dealer (and the dealer provides a quote or trades with). *No. CCY Pairs* counts the number of currency pairs in which customers request a quote from a dealer.

Summary Statistics

Panel A: Client (by Client Type)

	All (1705)	Corporation (1163)	Bank (419)	Institutional Investor (122)	Retail (1)
RFQ Volume	67.37	37.56	146.24	81.25	0.22
Trade Volume	44.89	33.66	66.97	76.50	0.22
No. RFQ	111.37	72.62	225.58	89.35	8.00
No. Trades	68.73	49.87	116.90	83.51	8.00
No. Dealers (Quote Requested)	6.39	6.10	7.61	5.02	1.00
No. Dealers (Quote Received)	5.54	5.50	6.08	4.16	1.00
No. Dealers (Traded)	3.3	3.12	3.90	2.91	1.00
No. CCY Pairs	6.22	5.21	8.79	7.07	1.00
No. CCY Pairs (Traded)	5.84	5.11	7.60	6.75	1.00
Market share		0.51	0.37	0.12	0.00

Panel B: Clients (by Client Size (Trade Volume))

	All (1705)	Top 10	11-30	31-75	76-190	191-500	501-1000	1001+
RFQ Volume	67.37	2,108.99	1,063.67	616.49	193.45	52.68	9.14	2.29
Trade Volume	44.89	1,646.06	729.16	339.93	129.93	35.13	7.64	0.78
No. RFQ	111.37	1,319.7	1,889.95	689.64	569.12	79.82	24.82	7.46
No. Trades	68.73	1,036.6	678.55	449.44	348.99	61.19	22.04	4.1
No. Dealers (Quote Requested)	6.39	28.3	18.55	15.11	11.9	9.4	5.7	3.45
No. Dealers (Quote Received)	5.54	27.1	17.1	12.56	10.53	8.16	4.9	2.95
No. Dealers (Traded)	3.3	24.0	15.7	10.71	7.79	4.85	2.55	1.3
No. CCY Pairs	6.22	23.8	32.25	24.76	15.96	9.31	4.66	2.2
No. CCY Pairs (Traded)	5.84	22.1	31.4	22.62	15.16	8.74	4.52	1.94
Corporation	0.68	0.6	0.4	0.38	0.61	0.65	0.72	0.71
Bank	0.25	0.2	0.45	0.56	0.32	0.28	0.22	0.21
Institutional Investor	0.07	0.2	0.15	0.07	0.07	0.07	0.06	0.08
Market share		0.22	0.19	0.2	0.2	0.14	0.05	0.01

Panel C: Dealer (by Dealer Size)

	All (334)	Top 5	6-15	16-50	51-100	101-200	201+
Volume RFQ (Quote Provided)	3,291.12	38,261.94	31,051.59	10,925.01	2,913.92	593.62	74.43
Trade Volume	229.16	3,400.42	2,084.79	657.4	212.82	46.99	2.54
No. RFQ (Quote Provided)	3,593.55	38,548.4	35,833.5	11,682.29	2,686.1	741.41	237.64
No. Trades	350.83	3,765.4	2,568.8	722.6	297.98	282.59	31.44
No. Clients (Quote Requested)	32.21	300.8	241.1	98.06	32.26	12.83	3.84
No. Clients (Quote Provided)	28.3	269.6	211.5	89.0	28.14	10.89	2.83
No. Clients (Traded)	16.84	189.0	133.8	49.46	16.6	5.9	1.42
No. CCY Pairs (Quote Requested)	36.86	141.2	126.1	92.34	52.22	29.06	11.9
No. CCY Pairs (Quote Provided)	28.84	133.6	112.7	73.43	41.66	22.95	6.64
No. CCY Pairs (Traded)	15.73	99.4	72.5	40.63	23.22	11.06	2.55
Market share		0.22	0.27	0.3	0.14	0.06	0

Table 1.2: Relationship between Client Network Size and Client Size: This table shows how the network size is related to client size. **Panel A:** shows how a client's size is related to the size of the client's network, measured by the number of dealers that a client requests a quote from. **Panel B:** shows how client size is related to the number of dealers that provide a quote to a client. **Panel C:** shows how client size is related to the number of dealers that a client trades with. *No. CCY Pairs* are the number of currency pairs in which a client is provided a quote. *Inst,Corp,Retail* are dummy variables equal to 1, if the client is an institutional investor, corporation or retail investor. Banks are the unreported client type.

Network Size by Client Size (for Different Network Measurements)			
Panel A: Number of Dealers (Quote Requested)			
	(1)	(2)	(3)
const	4.65***	4.05***	4.43***
	(0.17)	(0.19)	(0.34)
Client Size	1.28***	0.98***	0.97***
	(0.06)	(0.07)	(0.07)
No. CCY Pairs		0.16***	0.16***
		(0.02)	(0.02)
Inst			-1.8***
			(0.63)
Corp			-0.34
			(0.36)
Retail			-2.14
			(6.09)
N	1,660.0	1,660.0	1,660.0
R^2	0.2183	0.2498	0.2535

Panel B: Number of Dealers (Quote Received)			
	(1)	(2)	(3)
const	4.01***	3.43***	3.12***
	(0.15)	(0.16)	(0.29)
Client Size	1.14***	0.85***	0.84***
	(0.05)	(0.06)	(0.06)
No. CCY Pairs		0.16***	0.16***
		(0.02)	(0.02)
Inst			-1.19**
			(0.54)
Corp			0.53*
			(0.31)
Retail			-1.02
			(5.24)
N	1,660.0	1,660.0	1,660.0
R^2	0.2274	0.2664	0.2722

Panel C: Number of Dealers (Traded)			
	(1)	(2)	(3)
const	2.07***	1.42***	1.33***
	(0.1)	(0.09)	(0.17)
Client Size	0.92***	0.59***	0.58***
	(0.03)	(0.03)	(0.03)
No. CCY Pairs		0.18***	0.18***
		(0.01)	(0.01)
Inst			-0.44
			(0.32)
Corp			0.16
			(0.18)
Retail			0.37
			(3.1)
N	1,660.0	1,660.0	1,660.0
R^2	0.3254	0.4361	0.4376

Table 1.3: Transaction Cost by Client Size: This table shows how client size and network size are related to the transaction cost in an RFQ, focusing only on executed RFQs (trades). *Client Size* is the logarithm of the trade volume of a client. The network of the client is measured as *1/Dealers in RFQ*, i.e., the inverse of the number of dealers that provide a quote in the given client's RFQ; and as *1/Client Network Size*, the inverse of the total number of dealers that provide a quote to the client. *Client HHI* is the Herfindahl–Hirschman Index of the share of a client's trading by a dealer, that is, concentration in trading by clients. *Trade Size in mil*, *Log Trade Size in mil* are, respectively, the amount requested in mil EURO and its logarithm. *Institutional,Corporation* are dummy variables equal to 1, if the client is an institutional investor or corporation. Retail investors are subsumed by fixed effects, and banks are the unreported client type. *Buy & Sell Quote Provided* is a dummy variable equal to 1 if the RFQ requested quotes in both directions of the trade. All regressions include *Date, Time of Day* and *Currency Pair Fixed Effects*. Specification (5) introduces also *Dealer Fixed Effects*.

	Transaction Cost (by Client Size)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Client Size	-0.52***	-0.36***	-0.36***	-0.33***	-0.25***	-0.16***	-0.16***	-0.15***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
1/Dealers in RFQ		1.97***				1.9***		
		(0.02)				(0.02)		
1/Client Network Size			1.74***				2.13***	
			(0.02)				(0.03)	
Client HHI				1.93***				1.43***
				(0.02)				(0.03)
Trade Size in mil	0.12***	0.12***	0.11***	0.12***	0.09***	0.09***	0.09***	0.09***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Log Trade Size in mil	-0.12***	-0.04***	-0.05***	-0.06***	-0.05***	-0.03***	-0.04***	-0.04***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Inst	0.04*	0.14***	0.23***	0.17***	0.14***	0.02	0.12***	0.1***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Corp	-0.06***	-0.28***	-0.05***	-0.14***	-0.11***	-0.26***	-0.1***	-0.16***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Buy & Sell Quote Provided	0.15***	0.29***	0.21***	0.29***	0.13***	0.2***	0.18***	0.21***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
N	116,934	116,934	116,934	116,934	116,934	116,934	116,934	116,934
R^2	0.4275	0.4961	0.4765	0.4721	0.7284	0.746	0.7378	0.7343
Currency Pair F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time of Day F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dealer F.E.	No	No	No	No	Yes	Yes	Yes	Yes

Table 1.4: Transaction Cost by Client Size and Client's Average Amount of Trading with Dealers: This table shows how client size and network size are related to the transaction cost in an RFQ, focusing only on executed RFQ (trades). *Client Size* is the logarithm of the trade volume of a client. The network of the client is measured via *1/Dealers in RFQ*, i.e., the inverse of the number of dealers that provide a quote in the given client's RFQ; and *Log Client Network Size* is the logarithm of the total number of dealers that provide a quote to the client. Client's Average Trading with Dealers is measured via *Client Size/Client Network Size*, the client size divided by the number of dealers that a client receives quotes from, so $\frac{\text{Client Size}}{\text{Client Network Size}}$. *Trade Size in mil, Log Trade Size in mil* are the amount requested in mil EURO and its logarithm. *Institutional,Corporation* are dummy variables that are equal to 1 if the client is an institutional investor or corporation. Retail investors are subsumed by fixed effects, and banks are the unreported client type. *Buy & Sell Quote Provided* is a dummy variable equal to 1 if the RFQ requested quotes in both directions of the trade. All regressions include *Date, Time of Day* and *Currency Pair Fixed Effects*. Specifications (3) and (4) introduce also *Dealer Fixed Effects*.

Transaction Cost (by Client Size and Client's Average Trading with Dealers)				
	(1)	(2)	(3)	(4)
Client Size	-0.26***	0.05***	-0.05***	0.03***
	(0.0)	(0.01)	(0.0)	(0.0)
Log Dealers in RFQ	-0.95***		-0.69***	
	(0.01)		(0.01)	
Log Client Network Size		-1.69***		-0.88***
		(0.01)		(0.01)
Client Size/Client Network Size	-0.27***	-0.75***	-0.43***	-0.57***
	(0.01)	(0.01)	(0.01)	(0.01)
Trade Size in mil	0.12***	0.13***	0.08***	0.09***
	(0.0)	(0.0)	(0.0)	(0.0)
Log Trade Size in mil	-0.07***	-0.06***	-0.03***	-0.03***
	(0.0)	(0.0)	(0.0)	(0.0)
Institutional Investor	-0.27***	-0.32***	-0.21***	-0.13***
	(0.02)	(0.02)	(0.02)	(0.02)
Corporation	-0.42***	-0.19***	-0.29***	-0.12***
	(0.01)	(0.01)	(0.01)	(0.01)
Buy & Sell Quote Provided	0.36***	0.37***	0.15***	0.17***
	(0.02)	(0.02)	(0.01)	(0.01)
N	116,934	116,934	116,934	116,934
R^2	0.486	0.5061	0.7461	0.7427
Currency Pair F.E.	Yes	Yes	Yes	Yes
Date F.E.	Yes	Yes	Yes	Yes
Time of Day F.E.	Yes	Yes	Yes	Yes
Dealer F.E.	No	No	Yes	Yes

Table 1.5: Quoted Transaction Cost by Bilateral Trade Volume: This table shows how bilateral trading is related to the prices that are offered to clients by dealers. All specifications include RFQ fixed effects, a fixed effect for each trade direction of a RFQ a client initiates. The dependent variable is the half-spread in log bps quoted by a dealer to a client, and *Log Bilateral Trade Volume* is the logarithm of the trade volume between the client and the dealer offering the quote to the client. Specification (2) adds dealer fixed effects to the regression. Specifications (3) and (4) restrict the sample to RFQ and the trade directions of a two-way RFQ that do not lead to a trade.

Quoted Transaction Cost by Bilateral Trade Volume				
	(1)	(2)	(3)	(4)
Log Bilateral Trade Volume	-0.06***	-0.05***	-0.06***	-0.05***
	(0.0)	(0.0)	(0.0)	(0.0)
N	1,804,270	1,804,270	1,052,185	1,052,185
R^2	0.6477	0.6769	0.6437	0.6786
Within R^2	0.0249	0.0223	0.0296	0.0274
RFQ F.E.	Yes	Yes	Yes	Yes
Dealer F.E.	No	Yes	No	Yes

Table 1.6: Simultaneous Quote Provision to Multiple Clients: This table shows, for a given dealer, the price difference between the quote a larger client receives versus a smaller client. The sample is constructed by looking at dealers that provide the same quote to two different clients for the same trade size and same currency pair within 1 second of each other. Row *All* looks at all occurrences where both the smaller client and the larger client trade with the dealer at some point in time, row *Larger Client trades more* represents cases in which, in addition to trading with the same dealer as a smaller client, the larger client also trades more with that dealer than the smaller client, while row *Larger Client trades less* represents cases in which the smaller client trades more with the dealer. There are no cases in which the two clients are the same size or trade the same amount with the dealer, when converted to EUR. The columns are as follows: *N* is the number of observations used for each row, *Large No Quote* is the fraction of observations where only the large dealer does not receive a quote from the dealer, *<0* is the share of observations where the larger client receives a worse price, *=0* is the share of observations where the larger client and the smaller client receive the same price, while *>0* is the share of observations where the smaller client receives a better price than the larger client (always provided by the same dealer). Finally, *Small No Quote* is the fraction of observations where only the smaller client receives no quote from the dealer and *Both No Quote* is the fraction of observations, where both the smaller and larger clients receive no quote from the dealer, again all conditional on both clients trading with the dealer at some point in time.

Simulatenous Quote Provision to Multiple Clients (for larger Client)							
	N	Large No Quote	<0	=0	>0	Small No Quote	Both No Quote
All	245,887	1.9%	25.5%	30.1%	35.2%	4.5%	2.8%
Larger Client trades more	187,949	1.8%	18.2%	31.3%	41.3%	4.6%	2.8%
Larger Client trades less	57,938	2.1%	49.3%	26.3%	15.4%	3.9%	3.0%

Table 1.7: Quoted Dealer Spread (by Client Size and Bilateral Trade Volume): This table shows how bilateral trading and client size are related to the spreads offered by a dealer, when the dealer provides both a buy and a sell quote. The dependent variable is the spread in bps quoted by a dealer to a client based on the buy and sell quote provided by the dealer, and *Client* is the logarithm of the client's trade volume. *Log Bilateral Trade Volume* is the logarithm of the trade volume between the client and the dealer offering the quote to the client. *Trade Size in mil, Log Trade Size* is the (logarithm of the) trade size the client requests the quote for. *Inst, Corp* are dummy variables equal to 1 if the client is an institutional investor or corporation. Banks are the unreported client type, and no clients are retail investors. All specifications include dealer fixed effects, currency pair fixed effects, day fixed effects and time of day fixed effects. Specifications (3) and (4) introduce *Dealers in RFQ fixed effects* which is a fixed effect for the number of dealers in the RFQ. Specification (5) introduces client fixed effects.

Quoted Dealer Spread (by Client Size and Bilateral Trade Volume)					
	(1)	(2)	(3)	(4)	(5)
Client Size	-0.01***		0.03***		
	(0.0)		(0.0)		
Log Bilateral Trade Volume		-0.15***		-0.15***	-0.2***
		(0.0)		(0.0)	(0.0)
Trade Size in mil	0.24***	0.25***	0.24***	0.24***	0.22***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Log Trade Size	-0.13***	-0.13***	-0.13***	-0.11***	-0.09***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Institutional Investor	-0.18***	-0.05***	-0.26***	-0.04***	
	(0.01)	(0.01)	(0.01)	(0.01)	
Corporation	-0.02***	0	-0.09***	-0.04***	
	(0.01)	(0.01)	(0.01)	(0.01)	
N	564,292	564,292	564,292	564,292	564,292
R^2	0.6423	0.6484	0.6458	0.6519	0.6706
Dealer F.E.	Yes	Yes	Yes	Yes	Yes
Currency Pair F.E.	Yes	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes	Yes
Time of Day F.E.	Yes	Yes	Yes	Yes	Yes
Dealers in RFQ F.E.	No	No	Yes	Yes	Yes
Client F.E.	No	No	No	No	Yes

Table 1.8: Example of Riskless Principal Transaction: Panel A: shows an example of an RFQ where a match maker matches a client to another dealer via a riskless principal transaction (RPT). **Panel B:** shows an example where a match maker matches a client to other dealer that does not lead to a trade. Dates are reported in MM/DD/YY format.

Example of Riskless Principal Transaction								
Panel A: Riskless Principal Transaction (RPT) with trade								
Request ID	Tradetime	Quote Status	Requester	Dealer	CCY Pair	Buy/Sell	Notional	Quote
447915499	10/01/19 12:03:10.645	EXEC	Client 1	MM 1	EURUSD	Buy	300,000	1.08984
447915500	10/01/19 12:03:10.609	Not-Executed	MM 1	Dealer A	EURUSD	Buy	300,000	1.08980
447915500	10/01/19 12:03:10.609	Not-Executed	MM 1	Dealer B	EURUSD	Buy	300,000	1.08979
447915500	10/01/19 12:03:10.609	EXEC	MM 1	Dealer C	EURUSD	Buy	300,000	1.08979
447915500	10/01/19 12:03:10.609	Not-Executed	MM 1	Dealer D	EURUSD	Buy	300,000	1.08979
Panel B: Riskless Principal Transaction (RPT) without trade								
Request ID	Tradetime	Quote Status	Requester	Dealer	CCY Pair	Buy/Sell	Notional	Quote
447741069	10/01/19 07:45:51.109	Not-Executed	Client 2	MM 2	EURUSD	Buy	800,000	1.08862
447741070	10/01/19 07:45:51.124	Not-Executed	MM 2	Dealer E	EURUSD	Buy	800,000	1.08853
447741070	10/01/19 07:45:51.124	Not-Executed	MM 2	Dealer F	EURUSD	Buy	800,000	1.08850
447741070	10/01/19 07:45:51.124	Not-Executed	MM 2	Dealer G	EURUSD	Buy	800,000	1.08850
447741070	10/01/19 07:45:51.124	Not-Executed	MM 2	Dealer A	EURUSD	Buy	800,000	1.08851
447741070	10/01/19 07:45:51.124	Not-Executed	MM 2	Dealer H	EURUSD	Buy	800,000	1.08854

Table 1.9: Summary Statistics of Match Makers: This table reports the mean, standard deviation, and 25-, 50-, and 75-percentile of observations aggregated to the dealer level. All dealers are equally weighted. *All Match Makers* aggregates the statistic across all match makers and thus treats it as if all match makers are one. *Trade Volume* reports the volume the match maker trades acting as a dealer to a client *(Dealer)*, as a client requesting a quote *(Client)* or the trade volume intermediated through RPT. After *Trade Volume* all statistics are limited to the RPT via RFQ. *RPT No. Trades* are the number of RPT in RFQ that a match maker executes, while *Currency Pairs* are the currency pairs of those RPT in RFQ that the match maker is requested to provide a quote, while *Currency Pairs (Traded)* are the currency pairs that the match maker ultimately trades. *Avg. Trade Size (RPT) in* 10^5 is the average trade size of the trade size a match maker intermediates via RPT in RFQ. *No. Clients* are the number of clients that request a quote from a match maker, while *No. Clients (Traded)* are the number of those clients the match maker trades with. Finally, *No. Dealers* are the number of dealers the match maker receives a quote from, when using RPT, while *No. Dealers (Traded)* are the number of dealers the match maker trades with using RPT.

Summary Statistics of Match Makers

	N	μ	σ	25%	median	75%	All Match Maker
Trade Volume in RFQ (Dealer) in mil	46	80.42	129.24	4.24	14.81	72.28	3,699.21
Trade Volume in RFQ (Client) in mil	46	83.88	134.89	4.54	14.7	72.43	3,858.26
Trade Volume in RFQ (RPT) in mil	46	79.47	127.7	4.24	14.7	72.15	3,655.83
RPT No. Trades	46	132.26	220.27	6.25	32.5	163.5	6,084.0
Currency Pairs (RPT)	46	12.91	15.54	3.0	7.5	16.75	95.0
Currency Pairs (Traded) (RPT)	46	8.59	10.67	2.0	5.0	11.0	72.0
Avg. Trade Size (RPT) in 10^5	44	10.64	14.17	3.08	6.47	9.38	6.01
No. Clients	46	8.09	12.68	1.0	4.0	7.75	339.0
No. Clients (Traded)	46	4.65	6.3	1.0	2.0	5.0	202.0
No. Dealers (RPT)	46	6.91	6.48	2.25	5.0	8.75	84.0
No. Dealers (Traded) (RPT)	46	4.37	4.06	1.0	4.0	5.0	72.0

Table 1.10: Share of Exclusive Clients (of Match Maker) by Clients Size: This table shows how client size is related to trading exclusively with a match maker (being an exclusive client). Specifications (1) and (2) are OLS regressions, while Specifications (3) and (4) are Logit regressions. The independent variable is *Trades exclusively with a Match Maker*, an indicator variable that is equal to 1 if the client trades only with a match maker. *Client Size* is the log trade volume of a client, while *501-1000,201-500,81-200,31-80,11-30,Top 10* are indicator variables for the client being ranked within the corresponding group. The unreported group is *1001+*, which covers the 660 smallest clients (only 1660 out of 1705 clients trade). *Inst, Corp, Retail* are indicator variables equal to 1 if the client is an institutional investor, a corporation or a retail investor. Banks are the unreported client group.

Share of Exclusive Clients (of Match Maker) by Client Size

	(1 OLS)	(2 OLS)	(3 Logit)	(4 Logit)
const	0.1744***	0.1803***	-1.49***	-1.47***
	(0.0127)	(0.0145)	(0.15)	(0.19)
Client Size	-0.0048**		-0.07**	
	(0.0023)		(0.04)	
501-1000		-0.004		0.02
		(0.0143)		(0.24)
201-500		-0.011		-0.2
		(0.0168)		(0.29)
81-200		-0.0534**		-1.01*
		(0.024)		(0.54)
31-80		-0.0563		-0.62
		(0.0356)		(0.6)
11-30		-0.1084**		-0.77
		(0.0548)		(1.02)
Top 10		-0.075		-0.31
		(0.0768)		(1.46)
Inst	-0.1021***	-0.1022***	-1.08***	-1.31***
	(0.025)	(0.025)	(0.39)	(0.43)
Corp	-0.1382***	-0.1415***	-1.9***	-1.89***
	(0.014)	(0.0141)	(0.22)	(0.22)
Retail	-0.1816	-0.1803	-26.42	-0.07
	(0.2414)	(0.2413)	(1088981.22)	(2.63)
N	1,660.0	1,660.0	1,660	1,660
R^2	0.0562	0.0601	0.0593	0.0653

Table 1.11: Direct and Indirect Network Size for Match Maker Clients: This table shows how a match maker does not behave differently from a client in relation to her network size (the number of dealers that provide her with quotes). *Client Size* is the logarithm of trade volume and the measure of client size. *Match Maker* is an indicator variable for the client being a match maker, while *Contacts MM* is an indicator equal to 1 if the client contacts a match maker, while *Exclusive with MM* is an indicator for clients only contacting a single match maker. *No. CCY Pairs provided* are the number of currency pairs in which a client is provided a quote. *Inst,Corp,Retail* are dummy variables equal to 1 if the client is an institutional investor, corporation or retail investor. Banks are the unreported client type. Specifications (1)-(3) measure the number of dealers that a client contacts directly, while Specifications (4)-(5) measure the number of indirect dealers (a measure of the number of dealers a client contacts directly, not including the match maker, together with the number of additional dealer a match maker provides access to via RPT). For clients for whom the match maker never initiates an RPT the number of direct dealers is the same as the number of indirect dealers.

Direct and Indirect Network Size for Match Maker Clients

	(Direct, 1)	(Direct, 2)	(Direct, 3)	(Indirect, 4)	(Indirect, 5)
const	3.12***	3.11***	3.1***	3.39***	4.57***
	(0.29)	(0.3)	(0.31)	(0.33)	(0.36)
Client Size	0.84***	0.84***	0.84***	0.84***	0.91***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
Match Maker		0.16	0.17	-0.1	-1.28
		(0.82)	(0.83)	(0.88)	(0.93)
Contacts MM			0.02	6.44***	
			(0.41)	(0.43)	
Exclusive with MM					1.45**
					(0.62)
No. CCY Pairs provided	0.16***	0.16***	0.16***	0.17***	0.19***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Inst	-1.19**	-1.18**	-1.18**	-1.56***	-2.39***
	(0.54)	(0.55)	(0.55)	(0.58)	(0.62)
Corp	0.53*	0.55*	0.55*	0.2	-0.7*
	(0.31)	(0.32)	(0.33)	(0.34)	(0.37)
Retail	-1.02	-1.01	-1.01	-1.29	-2.38
	(5.24)	(5.24)	(5.24)	(5.54)	(5.89)
N	1,660.0	1,660.0	1,660.0	1,660.0	1,660.0
R^2	0.2722	0.2722	0.2722	0.3592	0.2752

Table 1.12: Transaction Cost when Trading with a Match Maker: This table shows the transaction cost that clients face. The dependent variable is *Transaction cost in bps*, the difference between the dealer's quote the client trades at and the midquote in log-bps ($\log(p_{i,t} - m_t) \times 10000$). *Executed with MM* is a dummy variable that is equal to 1 if the client trades with a match maker. This variable is further conditioned into either *Executed with MM x RPT* which is a dummy variable equal to 1 if the trade with the match maker is passed on via RPT, or *Executed with MM x non-RPT,* which is a dummy variable equal to 1 if the trade is not passed on via RPT and instead held by the match maker (for at least 1 second). *Exclusive with MM* is a dummy variable equal to 1 if a client exclusively receives quotes from a match maker, which is also further conditioned into *Exclusive with MM x RPT*, which is a dummy variable equal to 1 if the trade with the match maker is passed on via RPT, or *Exclusive with MM x non-RPT*, which is a dummy variable equal to 1 if the trade is not passed on via RPT and instead held by the match maker (for at least 1 second). The network size is controlled for by *1/No. Dealers in RFQ,1/Client Network Size* and *Client HHI*. *1/No. Dealers in RFQ* is the inverse of the number of dealers that provide a quote to the client in an RFQ, while *1/Client Network Size* is the number of dealers that provide quotes to the client at any point in time, and *Client HHI* is the Herfindahl–Hirschman Index of the market shares the client's dealers have with the client (measured in 0 to 1). *Client Size* is the log of the trade volume by a client. *Trade Size in mil, Log Trade Size* control for the trade size by the trade size in million or the logarithm of the trade size in million. Finally, *MM Executes* is a dummy variable equal to 1 if the RFQ is initiated by a match maker. All specifications have *Currency Pair, Time of Day* and *Day fixed effects* and *Dealer Group fixed effects*. The *Dealer Group*s are constructed as in Panel C of Table 1.1, so groups are *Top 5, 6-15, 16-50, 51-100, 101-200, 201+*.

Transaction Cost when Trading with a Match Maker

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Executed with MM x RPT	-0.06**	-0.39***	-0.44***	-0.38***			
	(0.02)	(0.02)	(0.02)	(0.02)			
Executed with MM x non-RPT	1.87***	1.66***	1.53***	1.66***			
	(0.06)	(0.06)	(0.06)	(0.06)			
Exclusive with MM x RPT					-0.4***	-0.39***	-0.31***
					(0.03)	(0.03)	(0.03)
Exclusive with MM x non-RPT					2.77***	2.67***	2.85***
					(0.07)	(0.07)	(0.07)
1/No. Dealers in RFQ		1.93***			1.92***		
		(0.02)			(0.02)		
1/Client Network Size			1.56***			1.54***	
			(0.02)			(0.02)	
Client HHI				1.8***			1.77***
				(0.02)			(0.02)
Client Size		-0.22***	-0.24***	-0.19***	-0.22***	-0.24***	-0.19***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Trade Size in mil	0.04***	0.07***	0.08***	0.08***	0.07***	0.08***	0.08***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Log Trade Size	-0.07***	0	-0.01***	-0.01***	0	-0.01***	-0.01***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
MM Executes	-0.31***	-0.11***	-0.21***	-0.21***	-0.11***	-0.2***	-0.2***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N	116,934	116,934	116,934	116,934	116,934	116,934	116,934
R^2	0.4885	0.5772	0.5579	0.559	0.5797	0.5601	0.5616
Currency Pair F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time of Day F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dealer Group F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Chapter 2

Centrality in OTC Markets, Liquidity Provision, and Prices

2.1 Introduction

In an over-the-counter (OTC) market, trading takes place in networks, in which a client forms relationships with a handful of dealers, and dealers, in turn, maintain relationships with multiple clients, as well as with other dealers (in the case of interdealer trading). Market participants' decisions on how many and which other participants to trade with determine their "position" in the network, formally captured by their (network) centrality. A substantial stream of research has sought to elucidate how the centrality of market participants affects pricing; such insights are critical for understanding price dispersion in OTC markets, as well as the broader functioning of these markets (Di Maggio et al., 2017; Hollifield et al., 2017; Li and Schürhoff, 2019; Hendershott et al., 2020a; Hasbrouck and Levich, 2021).

Thus far, studies in this literature tend to point to a centrality premium, in which the more-central party in a trade is the one that makes a profit. However, though their data identify which participant in a trade is the seller versus the buyer, these studies are limited by not observing which party initiates the trade—i.e., demands (vs. provides) liquidity—and thus cannot control for this factor. The importance of identifying liquidity demanders in studies of financial markets is well established, given that demanding liquidity is known to require the liquidity demander to pay a fee (Grossman and Miller, 1988). Indeed, much work has gone into accurately identifying the initiators of trades (i.e., liquidity demanders) in equity markets (Lee and Ready, 1991)—with Easley, de Prado and O'Hara (2016) accurately inferring trade initiation in modern equity markets. Yet, as will be elaborated in what follows, methods used for assigning trade initiation in equity markets are not sufficiently applicable to OTC markets—and thus, perhaps surprisingly, liquidity demanders (vs. providers) are typically not identified in these markets.

The current research overcomes the limitations of prior studies by using FX spot market trade data from a leading multi-dealer platform in the FX market, in which liquidity supply and demand are directly observed for each trade. To my knowledge, this study is the first to observe such information, and thus to be able to control for liquidity demand and provision. Analyzing the data, which cover 1.4 million request for quotes (RFQs) and Streaming trades that took place between late April 2019 and the end of 2019, I find that the centrality premium observed in previous studies can be attributed entirely to liquidity provision. That is, the liquidity demander (on average) always pays a fee to trade. More-central liquidity demanders pay lower fees compared with less-central liquidity demanders. Without controlling for liquidity demand and provision, I observe a centrality premium because the liquidity provider tends to be the more central party in a trade. However, controlling for liquidity demand and provision, I find a centrality discount, in which, for a given liquidity demander, transaction costs decrease as the liquidity provider's centrality increases.

I present the results in three steps. First, to establish comparability with prior studies, I characterize the relationship between centrality and trading costs in my data set while ignoring information on liquidity demand and provision. Specifically, I follow the paper closest to mine, Hasbrouck and Levich (2021), which also studies the FX market. Using the same information and regression setup as in that study, I find a centrality premium, similar in magnitude to that found by Hasbrouck and Levich (2021).

Then, in the second step, I study how my additional information (i.e., identification of the liquidity provider and demander in each trade) helps explain this centrality premium. Specifically, I describe the liquidity provision between two participants based on their centrality. I find that, in general, the more central participant tends to be the liquidity provider, whereas the less-central participant is the demander. Moreover, the greater the difference in centrality between the two participants in a trade, the greater the likelihood that the more-central participant is the liquidity provider.

In the third step, I control for liquidity provision and demand in each trade and show how the centrality of the liquidity demander and of the liquidity provider affect the transaction cost. I find that the liquidity demander pays a fee when trading, regardless of either participant's centrality. Thus, the centrality premium observed in the first step is entirely attributable to liquidity provision. At the same time, I observe that more-central liquidity demanders enjoy lower fees compared with less-central liquidity demanders. I further observe that the variation across liquidity demanders within a particular liquidity provider is much greater than the variation across liquidity providers but within a liquidity demander. Thus, it is primarily the centrality of the liquidity demander, rather than that of the liquidity provider, that determines the transaction costs.

In examining variation in transaction costs across liquidity providers, I find a centrality discount for peripheral liquidity demanders, i.e., trades with high-centrality liquidity providers (as compared with low-centrality providers) are associated with lower transaction costs. However, for more-central liquidity demanders, I find that the centrality of the liquidity provider does not affect the transaction cost the participant pays for demanding liquidity. In other words, for these liquidity demanders, transaction costs are similar across liquidity providers. I attribute these differences to the fact that more-central liquidity demanders have more trading partners than less-central (peripheral) liquidity demanders. Thus, while the most peripheral participants receive quotes from few liquidity providers (in many cases, just one), more-central participants receive quotes from many liquidity providers simultaneously; indeed, the average number of liquidity providers in an RFQ is 10. In the case of a more-central demander, a large number of quotes creates competition, in which the liquidity demander ultimately trades only with the liquidity provider supplying the best quote. The most competitive quotes happen to be similar across liquidity providers.

The latter finding highlights the fact that the transaction costs observed in my data are conditional on a trade taking place—that is, they reflect only the most competitive quote that a demander obtained, while ignoring all other quotes. Given that the average liquidity demander receives multiple quotes, I carry out a simulation to better understand how the results are influenced by conditioning on trades. I observe that if a given liquidity provider, on average, quotes better average prices than others, the ultimate differences in transaction costs will be much smaller than the difference in average quoted price. For example, with 10 liquidity providers, the differences in average quotes are passed through by less than 30% to the transaction cost; e.g., if a liquidity provider provides 0.05bps better quotes on average compared with her competitors, the transaction costs she realizes will only be 0.015bps lower than her competitors'.

On the other hand, an analysis of market shares corresponding to providers with different centrality reveals much greater differences. Continuing the previous example, the simulation suggests the liquidity provider who offers a better average quote of 0.05bps may attain a market share that is 70% larger than her competitors'. In other words, a better average quote means that a liquidity provider much more often supplies the best price, leading to a higher market share.

The empirical data supports a centrality discount conditional on liquidity demand, i.e., that more-central liquidity providers provide better average quotes. I find that a liquidity demander does not trade equally with all her liquidity providers. Instead, she trades significantly more frequently with her more-central liquidity providers. This finding indicates that the more-central liquidity providers, though not trading at better prices, provide better quotes on average, i.e., provide better quotes more frequently. Notably, these patterns are pronounced only among central liquidity demanders—who trade with many liquidity providers. Whereas peripheral liquidity demanders, whose networks are much smaller, tend to trade equally frequently with their liquidity providers, regardless of the provider's centrality. However, peripheral demanders trade at better prices with more-central providers.

The remainder of the paper is structured as follows. The following section provides a literature review, covering, among other aspects, centrality and transaction costs in OTC markets, as well as a discussion of the importance of identifying the initiator of a trade. Section 2.3 describes the data and institutional details. In Section 2.4, I construct my measure of centrality and carry out the first step of my analysis, showing the existence of a centrality premium when liquidity demand is not taken into account. Section 2.5 is the second step, where I show the patterns of liquidity provision and demand across market participants with different centrality. Section 2.6 presents the third step, showing how controlling for liquidity provision for each trade affects the observed centrality premium found in Section 2.4. In the subsequent Section 2.7, I provide a simulation to explain the results of Section 2.4, and Section 2.8 utilizes the results of this simulation to study market shares of liquidity providers. Finally, Section 2.9 concludes.

2.2 Literature Review

Previous research on network structure in OTC markets has consistently shown a core-periphery structure across all studied markets, including the corporate bond market (Di Maggio et al., 2017), the municipal bond market (Li and Schürhoff, 2019), the securitization markets (Hollifield et al., 2017), the Swap market (Riggs et al., 2020), and the FX market (Hasbrouck and Levich, 2021). A key question of interest is how transaction costs depend on the centrality of the participants involved. On the one hand, more-central participants may be more suited for intermediation and thus pass this advantage on to their trading partners through better prices, which would imply a centrality discount. On the other hand, central participants, given their high importance, may have high bargaining power, resulting in worse prices for their trading partners, which would result in a centrality premium.

Thus far, most studies of the relationship between centrality and prices have pointed to a centrality premium (Di Maggio et al., 2017; Li and Schürhoff, 2019; Hasbrouck and Levich, 2021). Nevertheless, Hollifield et al. (2017), who studied the securitization market, observed a centrality discount, and Hagströmer and Menkveld (2019), who also studied the FX market, found at most weak evidence of a centrality premium.

Notably, the studies outlined above do not provide a definitive explanation as to why a centrality premium or discount arises. Regarding the case of a centrality premium, Li and Schürhoff (2019), who studied the municipal bond market, argued that more-central liquidity providers provide quicker execution, but charge higher fees for this greater speed, compared with less-central providers. Hasbrouck and Levich (2021), in turn, provided some evidence that a centrality premium may arise due to liquidity provision by the central participants[1]. Hollifield et al. (2017) provided a model in which either a centrality discount or a centrality premium can arise, depending on the heterogeneity in liquidity demanders' sophistication. They argued that more-central liquidity providers charge higher transaction fees to each liquidity demander. However, they tend to trade more with more-sophisticated liquidity demanders, which receive lower transaction costs compared with less-sophisticated liquidity demanders.

All the works cited above observed the majority if not all trading in their respective markets, and dealer identifiers enabled them to study trading outcomes in interdealer trades[2] (Di Maggio et al., 2017) or in trades between dealers and clients (Hollifield et al., 2017; Li and Schürhoff, 2019). While studies on interdealer trades have the benefit of identifying both participants in the trade, studies on dealer-client trades identify the spread that each dealer charges the clients that she trades with[3]. However, none of these studies identified the party responsible for liquidity provision versus liquidity demand, and thus could not control for this factor—often resorting to the perspective of the buyer or the seller to evaluate transaction cost. In my study, I directly observe liquidity demand and provision, enabling me to gain insights regarding what drives

[1]The studied markets are very different, and thus each explanation may be applicable only to the respective market. Furthermore, in the illiquid municipal bond market, being able to trade immediately is a form of liquidity provision, something that is commonly available in the FX market.

[2]Hasbrouck and Levich (2021) identifies all participants, both clients and dealers, but, like other studies in this category, does not observe liquidity demand or provision.

[3]It is generally thought that a client contacts a dealer to trade. The dealer then may wait or actively search for a client to make the reverse trade. Studies in the latter category cannot identify which of the clients initiated the trade. Furthermore, they cannot determine whether any of the two clients received a price better or worse than the fundamental value of the asset, i.e., whether the second client was providing liquidity.

the centrality premium observed in the previous literature—and, more broadly, to characterize the relationship between centrality and the likelihood of providing versus demanding liquidity.

As noted above, Grossman and Miller (1988) showed that demanding liquidity generally requires the liquidity demander to pay a fee. On the basis of this idea, Lee and Ready (1991) developed an algorithm to assign liquidity demand to either the buyer or the seller in a trade in an equity market. Specifically, this algorithm assumes that if the price is above the midquote, the buyer is likely to be the liquidity demander. This approach has recently been shown to attain high accuracy in modern markets Easley et al. (2016). No such algorithms exist for OTC markets. First, it is problematic to apply these algorithms in OTC markets because many of these markets, e.g., the corporate bond and municipal bond markets, are illiquid, making it difficult to determine the midquote of the assets (Di Maggio et al., 2017). Second, even in liquid markets, like the FX market, where spreads are much smaller, minor discrepancies between the reporting time and trade time may result in the incorrect assignment of liquidity demand to a participant (Hasbrouck and Levich, 2021).

In cases in which the identities of dealers and clients are revealed, it is possible, to some extent, to identify liquidity demanders in OTC markets on the basis of the assumption that clients demand liquidity from dealers rather than vice versa. This approach is insufficient, however, given that interdealer trading is common in OTC markets, such that liquidity providers (dealers) are also demanders, selecting a set of liquidity providers that they demand liquidity from (Di Maggio et al., 2017; Hollifield et al., 2017). The interdealer market can account for substantial trading volume; for example, in the FX market, 30% of all trading volume is interdealer trading[4] (Bank of International Settlement, 2019). In these interdealer trades, as well as in riskless principal transactions, the participant initiating the trade is not identified.

In studies of OTC markets—in which most participants may choose whom they demand liquidity from, but themselves do not provide liquidity—the inability to directly observe liquidity demand and supply limits the capacity to characterize market dynamics in general, and to elucidate the relationship between centrality and transaction costs specifically, as liquidity demand is a major influence on the transaction cost. The current study overcomes this limitation and observes liquidity demand and provision directly, thereby providing several new insights. For example, as discussed in later sections, the observation of liquidity demand allows me to reconcile the conflicting results of Hasbrouck and Levich (2021), who find a centrality premium, and Hagströmer and Menkveld (2019), who find at best weak support for a centrality premium.

Hendershott, Li, Livdan and Schürhoff (2020b) showed that, in illiquid markets, reliance on trade data alone leads to an underreporting of transaction costs, since many requests fail to trade. I show that a similar problem arises in liquid markets, as participants receive many quotes simultaneously. A trade observes only the most competitive quotes by a liquidity provider, underestimating the differences between liquidity providers. In section 2.8, I show

[4] Note that 30% interdealer trading means that almost all trades with a liquidity demander require trading between two liquidity providers, as intermediating between liquidity demanders requires two trades, one for each side of the transaction.

how examining liquidity providers' market shares, rather than transaction costs, can help to provide a more accurate picture of differences across liquidity providers.

In addition to providing a more detailed perspective on the relationship between centrality and transaction costs, this study sheds light on the broader trading dynamics of central versus peripheral participants. Bernhardt et al. (2004) showed that, in interdealer trading, larger liquidity providers—who are also likely to be more central—receive better prices. At the same time, peripheral liquidity providers rely more on interdealer trading than do more-central dealers (Hollifield et al., 2017). Given these dynamics, Skiera (2021b) shows why liquidity demanders may still want to trade with peripheral (rather than core) liquidity providers. Specifically, some peripheral liquidity providers serve as intermediaries to core liquidity providers that the peripheral demander can only access at worse prices—and thus provide prices that are not available to these liquidity demanders. In other words, some peripheral liquidity providers pass on the better prices of core liquidity providers to liquidity demanders.

These insights notwithstanding, much remains to be elucidated regarding the overall dynamics of and reliance on interdealer trading by liquidity providers. The current study contributes such insights by observing not just the liquidity provision by many peripheral liquidity providers but also their liquidity demand.

This work further contributes to our understanding of electronic trading and the trading mechanisms arising from it. O'Hara and Zhou (2021) showed that electronic trading has steadily increased over the last decade even in illiquid markets; in liquid markets, like the FX market, trading is already predominantly electronic. Electronic platforms enable traders to easily contact multiple liquidity providers simultaneously, as opposed to having to telephone each prospective liquidity provider individually (Vogel, 2019).

The first trading mechanism to emerge on electronic platforms was an auction mechanism, namely RFQ; the current work studies this mechanism. Prior works examining RFQs include Hendershott and Madhavan (2015), who explored when RFQs are used, and Riggs et al. (2020), who studied competition in RFQs.

More recently, new electronic trading mechanisms have arisen that lead OTC markets to more closely resemble their competing market: the exchange. One such mechanism, which this work also studies, is Streaming, where liquidity providers continuously provide liquidity demanders executable quotes, resulting in a personalized limit order book to each liquidity demander. Riggs et al. (2020) provide first statistics on this new trading mechanism, but economic differences between Streaming and other trading mechanisms are still unexplored. This study provides several insights into the differences between Streaming and RFQ, in terms of the transaction costs realized by participants, and the relationships between those costs and participants' centrality.

Finally, this work relates to the microstructure of FX markets (see Evans and Rime (2019) for a recent literature review; see also the subsequent section for a description of the FX market structure). Most papers in this literature focus on interdealer trading in FX markets, as the continuous central limit order book is ideal for the study of price discovery (Evans and Lyons, 2002; Moore et al., 2016), and Hagströmer and Menkveld (2019) show that information diffuses from the central limit order book to the private quotes of dealers. Relatively few studies have explored the dealer-to-client market, as the highly fragmented trading environment makes it difficult to obtain records with multiple dealers (Menkhoff et al., 2016). However, Hasbrouck and Levich (2019) and Hasbrouck and Levich (2021), using CLS settlement data, identified settlement banks, and were the first to study the network structure of the FX market. The current study similarly overcomes data limitations and sheds light on liquidity provision and its effects on transaction costs in the FX market.

2.3 Institutional Background and Data Description

2.3.1 Foreign Exchange Market Structure

The FX (spot) market is a heavily fragmented OTC market with fragmentation across and within market segments. The market comprises two segments: the interdealer market and the dealer-to-client market. The interdealer market covers trading between dealers and includes centralized options; in this market, trading takes place electronically and uses one of two limit order books (Hasbrouck and Levich, 2021): EBS (e.g., (Hagströmer and Menkveld, 2019)) and Reuters, besides bilateral trading. Each currency pair predominantly trades in a single limit order book, with EBS focused on EUR, JPY, and CHF; and Reuters being the dominant trading platform for GBP, AUD, CAD, and the Scandinavian currencies (King et al., 2013).

The interdealer market, which covers 30% of trading (Bank of International Settlement, 2019), plays a key role in price discovery (Moore et al., 2016; Hagströmer and Menkveld, 2019). In particular, Hagströmer and Menkveld (2019) found that the interdealer market (EBS in their study) first incorporates price updates, which later lead to updates in the prices quoted by dealers. At the same time, Evans and Lyons (2002) showed that prices in the FX market are heavily influenced by client order flow, finding R^2 of above 60%, so that price discovery is very much about the demand for currencies by clients. While these limit order books are accessible almost exclusively to dealers, a minority of interdealer trading takes place in limit order books, with a much greater share taking place in OTC markets (Schrimpf and Sushko, 2019).

As clients generally do not have access to the interdealer market, their trading occurs in the dealer-to-client market; as this is an OTC market, trading is non-anonymous. In contrast to many other OTC markets, trading in the dealer-to-client FX market is mainly electronic (71% of spot volume) (Bank of International Settlement, 2019). The dealer-to-client market is much more fragmented compared with the interdealer market, and is also much less studied.

The dominant form of client-dealer trading is through single- and multi-bank platforms (Bjønnes and Kathiziotis, 2016; Bank of International Settlement, 2019). These are electronic platforms where dealers either respond to clients' RFQs or stream prices to their clients. The difference between single-bank and multi-bank platforms is that, in single-bank platforms, only one dealer (and maybe the dealer's prime brokerage clients) provides prices and thus liquidity to clients. In contrast, in multi-bank platforms, multiple dealers provide prices to clients. Bjønnes and Kathiziotis (2016) show that dealers are active on both types of platforms and that multi-bank platforms are the more common form of trading, executing 39% of all trading volume versus 12% on single-bank platforms. My data set covers trading from a leading multi-bank platform.

2.3.2 Trading on a Multi-Bank Platform

Trading on a multi-bank platform is non-anonymous, allowing dealers—in the role of liquidity providers—to personalize pricing, i.e., the provision of quotes, to each liquidity demander (i.e., clients or other dealers, in the case of interdealer trading). Before submitting an RFQ for a specified volume and currency pair, the liquidity demander decides which liquidity providers to contact and can contact any liquidity provider with whom the liquidity demander has a relationship. After selecting the liquidity providers to contact and determining the trade size and currency pair, the liquidity demander decides to either request both a sell and a buy price or request a price only in the direction of the desired trade. The default option is the latter, so the vast majority of RFQs only receive quotes in the liquidity demander's trade direction.

After receiving an RFQ from a liquidity demander, the liquidity provider decides whether to submit a quote and, if so, what quote to submit. Due to the fast-changing nature of the FX market and the tight pricing, liquidity providers have the option to change their price up to the time of execution by the liquidity demander. A liquidity demander selects at most one quote to trade, after which the dealer whose quote was selected has the right to reject the trade ("last look"). Though popular in other trading mechanisms, "last look" is uncommon in RFQ, as the liquidity providers first observe the request by the liquidity demander before providing quotes to the liquidity demander.

In Streaming, liquidity providers continuously provide the liquidity demander personalized price-quantity quotes to trade. The quotes by different liquidity providers then get aggregated into a liquidity demander-specific limit order book, where the liquidity demander has an overview of all the quotes and associated quantities available to her at each price. The liquidity demander then submits an order executing against the best quotes provided to her. If the order quantity exceeds the amount provided by a single quote at the best price, the order may execute against many quotes and at multiple prices, just like a market order in a limit order book.

Liquidity providers do not know when a liquidity demander trades with another liquidity provider and instead only observe against which of their quotes a liquidity demander wants to trade. As in an RFQ, the liquidity providers whose quotes were selected have the right to reject the trade ("last look"). Because in Streaming, unlike in RFQ, the liquidity provider provides the quotes continuously (rather than upon request), "last look" is more common in Streaming vs. RFQ (4.6% vs. <1.9% of trades) but remains uncommon. In either mechanism, after a liquidity provider evokes her "last look", the client most frequently trades with another liquidity provider.

2.3.3 Data Description

The data were provided by a leading multi-bank platform in the FX market, with the majority of trading being dealer-to-client trading. The platform is one of only six multi-bank platforms named in the Bank of International Settlement's discussion of the "Triennial Central Bank Survey of Foreign Exchange and Over-the-counter (OTC) Derivatives Markets in 2019" (Bank of International Settlement, 2019) in Schrimpf and Sushko (2019). Further anecdotal evidence supporting the platform's leading position in the market is reflected in the fact that, according to my discussions with platform managers, Citi—the third-largest dealer in the FX market (Euromoney, 2019)—continued to provide prices to this platform even after deciding, in late 2019, to reduce the number of platforms it provides pricing to by two-thirds (Financial Times, 2019).

As is typical in the FX market, trading on the platform is entirely electronic. Since the data is provided by one platform, the data set on major currency pairs used in this study can be thought of as constituting a subset of the settlement data studied in previous literature (Hasbrouck and Levich, 2021), but provides greater detail on each trade. Unlike data used in previous research (Hasbrouck and Levich, 2021; Li and Schürhoff, 2019; Di Maggio et al., 2017), my data set enables me not just to identify both participants in each trade, but also to observe which of the two participants demands liquidity and which participant provides it. Specifically, in an RFQ, the liquidity demander initiates the RFQ by requesting quotes from liquidity providers, while in Streaming, the liquidity demander decides to execute a quote provided to her by a liquidity provider.

The data set covers trading among 2,411 participants (2024 clients, participants that solely request quotes; and 387 dealers, which must provide liquidity, but can also demand liquidity) from April 25, 2019 to December 31, 2019, via RFQ and Streaming. The data set covers trading in 184 currency pairs —a much greater number of currency pairs relative to previous studies. All these currency pairs were traded via RFQ, and 112 of these currency pairs also traded via Streaming. See the appendix (Appendix B.1) for a detailed description of the sample construction.

The data set covers 637,816 RFQ orders, of which 404,065 orders led to trades. The relatively low conversion of RFQ orders to trades is due in part to dealers called "match makers" (Skiera, 2021b). These participants are dealers who, upon being contacted by a client, initiate an identical RFQ with another dealer, and trade with that dealer if and only if their client trades with them. Thus, the large number of unexecuted RFQs is largely attributable to match makers who are contacted by clients that ultimately decide not to trade with them. (See Skiera (2021b) for a detailed description of how match makers operate.) Controlling for the existence of match makers, an RFQ leads to trade in 94% of cases[5].

A similar picture emerges in Streaming, where the data set covers 962,041 Streaming orders, of which 927,602 lead to execution(s)—an execution rate of 96%, very similar to that of RFQ. Unlike in RFQ, an order in Streaming can execute against multiple quotes provided by liquidity providers, so the number of trades is larger than the number of executed orders (996,531 vs. 927,602). However, the average order leads to only 1.07 trades, and only 2.8% of orders execute against multiple quotes. Furthermore, conditional on executing an order, the fill rate is very high at 99%.

Table 2.1 shows the distribution of order and trade size, as well as transaction cost (*Half-Spread*), for RFQ and Streaming. For both RFQ and Streaming, the sizes of executed orders are similar to those of unexecuted orders, with almost no differences across the whole distribution. In Streaming, executed orders also execute for the amount that is requested, except for the largest orders, which achieve a 90% fill rate (*Executed Orders: Order Size* vs. *Executed Orders: Filled Size*). There are differences in order size between RFQ and Streaming. Specifically, though order sizes are similar for much of the distribution, beyond the 75-percentile, RFQ orders are larger than Streaming orders. For example, at the 90-percentile, RFQ orders are twice as large as Streaming orders, with an order size of 3.61mil EUR (vs. 1.85 in Streaming), and orders stay 1.2mil EUR larger in RFQ at the 95-percentile.

Given that Streaming allows for multiple trades per order, the trade size in Streaming is slightly smaller than the order size. However, the difference is slight, at only 60,000 EUR on average, since only large orders get split into multiple trades. In fact, below the 50-percentile, the trade size is even larger than the order size, whereas at the 95-percentile, the trade size is only two-thirds of the order size. This shift in distribution happens because large orders split and execute into multiple trades, with each of these trades generally having a trade size of 1 mil.

Between Streaming and RFQ, there are apparent differences in transaction costs (*Half-Spread*), with Streaming having significantly lower transaction costs (0.31bps vs. 0.75bps). This difference in transaction cost is due to four reasons: 1) Clients in Streaming trade at higher volumes compared with clients in RFQ; indeed, only the largest clients trade in Streaming, while all other clients trade in RFQ. 2) Clients in Streaming are more sophisticated than those in RFQ; specifically, Streaming clients include liquidity providers, institutional investors, and smaller banks, whereas corporations mainly use RFQ. 3) Streaming offers trading in a slightly smaller number of currency pairs, and excludes some of the most illiquid currency pairs. 4) Order sizes in Streaming are smaller and more concentrated around the round lot size of 1mil EUR.

[5]This fraction is calculated excluding the liquidity demand of banks, as many of these are match makers, if they demand liquidity. The fraction is thus calculated as the share of volume that corporations, institutional investors, or retail investors request that ends up being traded.

2.3.4 Order Flow Observation in the Data

Given that the data come from one platform in the FX market, they cover only part of all trading in this market. Because the FX market lacks a central repository, no researchers thus far have been able to observe all trading in the FX market (Hagströmer and Menkveld, 2019; Hasbrouck and Levich, 2021; Menkhoff et al., 2016), with Hasbrouck and Levich (2021) covering the greatest amount of trading. Though central repositories introduce certain disadvantages—specifically, in other OTC markets, such repositories only identify trades and do not reveal which party in a trade is the liquidity provider versus demander—the capacity to observe only partial trading may nevertheless create certain biases. In what follows, I will discuss how the partial observation of data may affect my results. This discussion addresses liquidity demanders and liquidity providers separately.

In general, many liquidity providers in FX markets utilize multiple trading venues to trade; for example, a prominent dealer might have its own single-bank platform, may provide quotes on a multi-bank platform, and access the interdealer limit order books. This utilization of multiple platforms is especially likely to occur among liquidity providers that do not demand liquidity on the platform, since even the largest liquidity providers must demand liquidity some of the time. And indeed, my data set does not observe the largest liquidity providers (around 40 participants) demanding liquidity and thus cannot characterize their demand patterns.

However, I expect the vast majority of liquidity demanders to solely use this platform for trading. The platform provides access to all the largest liquidity providers in the market. Liquidity demanders likely do not contact the same liquidity provider on multiple platforms, and the largest liquidity demanders contact up to 40 different liquidity providers on the platform. Thus, these up to 40 liquidity providers are likely their complete network. Furthermore, from the perspective of the liquidity demander, usage of the platform is free; trading fees are only levied on the liquidity provider, and the platform also provides many tools to integrate order execution into clients' other systems[6]. As integrating different systems is generally considered expensive, I expect small clients to exclusively use one platform.

In the case of participants that both demand and provide liquidity, it is possible to verify more concretely that these users trade exclusively on the platform. Specifically, of these 140 participants, only 4 provide more liquidity than they demand in the spot market, leading me to conclude that the others trade exclusively on the platform. Appendix B.2 provides a detailed description of these liquidity providers.

Given that the data set is likely to contain almost all trading by liquidity demanders, any bias caused by incomplete trading data is unlikely to impair the capacity of this research to address its primary focus—namely, characterizing pricing provided to different liquidity demanders, and understanding how prices differ across different liquidity providers, based on the parties'

[6]Besides the opportunity to trade, the multi-bank platform providing the data also provides the client with many additional services, such as documentation on regulatory compliance, such as MiFID II, pre- and post-trade analytics, as well as integration into enterprise management systems.

centrality[7].

2.4 Centrality and Trade Prices

As noted in the previous section, the sample consists of 2,411 participants. Of these 2,411 participants, 2,024 (84%) are solely liquidity demanders (i.e., participants that solely request quotes), and 387 (16%) are liquidity providers, participants that provide a quote at least once. Of these 387 liquidity providers, 140 (36%) not only provide quotes but also demand quotes from other liquidity providers on the platform[8]. These participants form 15,276 relationships in which quotes are provided, of which 10,438 lead to trades. Given the vast numbers of both clients and dealers—who collectively engage in more than 1 million trades—each participant forms a network in which to trade, as is common in OTC markets (Di Maggio et al., 2017; Li and Schürhoff, 2019; Hasbrouck and Levich, 2021).

Recall that the focus of this study is on the relationship between participants' centrality—i.e., their "position" in the network, as reflected in their centrality—and the transaction costs that these participants realize. To ensure comparability with previous studies of centrality and transaction costs in OTC markets, and particularly with the work of Hasbrouck and Levich (2021)—given that my data set can be thought of as a subset of theirs (though their data do not capture details on liquidity provision and demand)—I follow the previous literature closely. In this section, I show that my data replicate the findings of Hasbrouck and Levich (2021) when ignoring my additional information on liquidity provision and demand. In subsequent sections, I proceed to utilize these details to provide new insights into the relationship between centrality and transaction cost.

2.4.1 Network Construction

My measure of centrality is degree centrality, which Hasbrouck and Levich (2021) also use. A participant's degree centrality is defined as the number of trading partners a participant has. For liquidity demanders, the degree centrality is the number of liquidity providers with whom the liquidity demander trades. For liquidity providers, in turn, the degree centrality is the size of the set of liquidity demanders and liquidity providers with whom the liquidity provider trades.

[7]Concretely, I take the quotes provided by liquidity providers to liquidity demanders as given and do not assess whether the provided quotes are optimal for the respective liquidity provider, instead of focusing on the implications of these provided quotes to the liquidity demander and the network structure more generally.

[8]See Appendix B.2 for a detailed discussion of how liquidity providers are matched to liquidity demanders, and which liquidity providers are also liquidity demanders on the platform. This matching was done by hand and I decided to only include a matching if a high degree of certainty for the liquidity demander and liquidity provider being the same identity was given. As such, some matches may be missing.

Panel A of Table 2.2 reports the distribution of degree centrality on the platform. The average participant trades with 8.7 other participants over the 8-month period. However, the distribution is very skewed, as 50% of participants trade with just 3 participants, and even at the 90-percentile, the participants trade with just 16 participants. In total, 80% of participants trade with less than the average number of trading partners (8.7). Instead, the 10 participants with the most trading partners (Top 10) each trade with at least 250 participants and up to 449 participants, trading on average with 330.4 participants. The following 30 participants with the greatest number of trading partners (11-40) each trade with 137.6 participants on average. Outside the Top 115 (Top 5.7%) most central participants, the participants trade with at most 27 other participants, showing how skewed the distribution is.

To analyze how the centrality of a participant affects the realized trade prices, I sort participants into groups based on their degree centrality[9]. Following Hasbrouck and Levich (2021), I define the four most central groups such that each group is involved in 40% of all trade volume. (As trades have two sides, the involvement in trades adds up to 200%.) To provide insights regarding how transaction costs vary across more-peripheral participants, I split the last 40% of trading into three further groups. Of these three groups, the first group covers the remaining participants with a degree centrality larger than the mean and covering 20% of all trading. Each of the final two groups covers 10% of all trading. Panel A shows the groups and their sizes. The first group contains only 10 participants, while the following two groups contain 30 and 75 participants, respectively. In total, the Top 250 (10%) participants (in terms of centrality) account for 80% of trading.

Panel B of Table 2.2 shows that trading volume follows a similar distribution. The median participant trades 32.9mil EUR over the 8-month period, while the average participant trades 1.33bn EUR over the same period. Top 10 participants trade more than 5bn per month ($\frac{47.319}{8}$), far more than the average participant in the second group (11-40), who trades less than half that per month ($\frac{20.82}{8}$). At the same time, both the second and third groups contain very large participants that trade more than the smallest participant in the Top 10 group. Even the most peripheral group (1000+) has a participant that trades more than 1bn EUR per month[10]; however, the median trade volume in that group is just 1mil EUR a month.

2.4.2 Profit to Buyer by Centrality

In the absence of the capacity to observe or reliably infer trade initiation in OTC markets (see discussion above), but given data on the identity of the buyer and the seller, previous studies (Di Maggio et al., 2017; Hasbrouck and Levich, 2021) have examined the profit or loss to a particular side of the transaction, for example, the buyer. I follow Hasbrouck and Levich (2021)

[9]Participants, both clients and dealers, are sorted by their degree centrality together. In cases where two participants have the same degree centrality, the greater trade volume serves as the tiebreaker.

[10]The large volumes in group 1000+ are due exclusively to participants that decide to trade with 1 participant only and not participants trading with 2 to 4 participants.

by studying the profit to the buyer for each trade. The *profit to buyer* is defined as:

$$\textit{profit to buyer}_{i,j,t,c} = \left(\log\left(m_{t,c}\right) - \log\left(q_{i,j,t,c}\right)\right) \times 10,000 \tag{2.1}$$

with $m_{t,c}$ being the midquote and $q_{i,j,t,c}$ being the price at which the buyer i buys the quote currency for the base currency of currency pair c at time t from seller j. I report *profit to buyer* in log-bps, which, given that transaction cost is around only 1 log-bps, is indistinguishable from bps.

Table 2.3 reports the *profit to buyer* when a buyer in group *row* buys from a seller in group *column*. First, the upper-right triangle of Table 2.3 reports almost exclusively positive coefficients, while the lower-left triangle reports exclusively negative coefficients. This pattern means that the buyer makes a profit whenever a more-central buyer trades with a less-central seller (upper-right triangle). If the buyer is less central than the seller (lower-left triangle), the buyer makes a loss. Thus, the more central party makes an immediate profit on the transaction by acquiring the asset below the midquote price on average.

Second, the magnitudes of profits increase in the difference between the buyer's and seller's centrality. For example, buyers in the 11-40 group (in terms of their centrality ranking) pay 0.12bps to Top 10 sellers when trading with Top 10 sellers, but gain an insignificant 0.06bps when buying from another 11-40 participant. When buying from participants in either the 41-115 or the 116-250 groups, they gain around 0.3bps. Participants in the 11-40 group further gain 0.39bps when buying from 251-450 participants; 0.43bps when buying from the 451-1000 group; and 0.53bps when buying from the 1000+ group. Thus, in a trade, the more central party (buyer/seller) makes a profit, and the difference between the buyer's and seller's centrality determines the size of the profit the more central participant makes.

Third, Table 2.3 is roughly symmetric around the diagonal. So, while an 11-40 buyer makes a 0.31bps profit when buying from a 41-115 seller, a 41-115 seller makes a 0.34bps loss when buying from an 11-40 seller. Thus, a buyer/seller does not automatically make a profit. Table 2.3 rather captures something about trading between 11-40 and 41-115 participants, as a 41-115 participant on average makes a 0.3bps loss when trading with an 11-40 participant, regardless of whether the 41-115 participant is buying or selling.

These findings replicate the main finding in Hasbrouck and Levich (2021), showing that a centrality premium exists, i.e., the more central party makes a profit on the trade. The magnitudes of the centrality premium are greater in this study than in Hasbrouck and Levich (2021). The difference arises from the construction of the centrality groups, the greater focus on trading with more-peripheral participants, and trading in many more (and less liquid) currency pairs. The three most central groups in Hasbrouck and Levich (2021), which are constructed at the currency pair level, generally have just 5 participants each (an additional restriction in the construction), while the fourth most central group often has around 10-20 participants, containing the 16-36 most-central participants. Thus, group 11-40 in this paper likely represents Hasbrouck and Levich (2021)'s fourth most central group. As a comparison, Hasbrouck and

Levich (2021) report that their most central group makes a profit of 0.10bps when trading with the fourth most central group. In this study, a Top 10 participant makes a profit of 0.12-0.17 bps when trading with an 11-40 participant (the corresponding comparison group). The observed centrality premium for a particular central participant in this paper is very comparable to that in Hasbrouck and Levich (2021), especially when accounting for the inclusion of currency pairs that are less liquid, and thus associated with higher transaction costs.

2.5 Centrality and Liquidity Provision

The findings elaborated above, which identified a centrality premium in the FX market (when liquidity demand and provision were not taken into account), are in line not only with those of Hasbrouck and Levich (2021) but also with the findings of other studies of OTC markets—for example, the works of Di Maggio et al. (2017) and Li and Schürhoff (2019), who studied the network structures of the corporate bond and municipal bond markets, respectively. Still, it is unclear why a centrality premium exists. Hasbrouck and Levich (2021) looked at the direction of flows between more-central and less-central participants and found that central participants buy at times when the currency pair appreciates, consistent with liquidity provision by more-central participants. In this section, I utilize the capacity to identify the liquidity provider and demander—a feature unique to my data set, as discussed in previous sections—to explain the centrality premium observed in the FX market.

2.5.1 Liquidity Provision between Centrality Groups

Table 2.4 shows, for each pair of centrality groups, the share of trades in which the less-central group demands liquidity from the more-central group. The results are striking, indicating that, in almost all cases, the less-central participant requests liquidity from the more-central participant (and conversely, the more-central participant generally provides liquidity to the less-central participant). Furthermore, the share of liquidity provision by the more central group generally increases with the difference in centrality between the two groups. Table 2.4 thus suggests that the centrality premium observed in Table 2.3 is influenced by the liquidity provision, as the liquidity provider is expected to be paid a fee for providing liquidity to the liquidity demander.

The full extent of how liquidity provision affects the observed centrality premium will be studied in Section 2.6; however, a first hint regarding the importance of liquidity provision can be seen from the one exception in Table 2.4 where the less-central participant provides more liquidity to the more central participant. Specifically, when examining the centrality groups 451-1000 and 116-250, I observe that the higher-centrality group (116-250) demands liquidity in 69% of trades. Furthermore, in Table 2.3, 116-250 buying from 451-1000 is the one cell in the upper triangle where the more-central buyer actually makes a loss when trading. This finding suggests that liquidity provision may be the driving factor behind the centrality premium, and that the centrality premium may simply reflect the fees that liquidity demanders pay to liquidity providers—who tend to be more central than the demanders.

2.5.2 Liquidity Provision by Centrality Group

Figure 2.1 shows, for each centrality group, the share of trades in which the participant in that group demanded versus provided liquidity. The two most-central groups (Top 10 and 11-40) provide liquidity in almost all their trades; this observation is the result of large liquidity providers not demanding liquidity on the platform. For all other groups, however, I expect to observe all trading for any participant demanding liquidity (as discussed in Section 2.3), so that Figure 2.1 reveals complete information about these participants' liquidity demand. For these groups, the share of liquidity provision is dramatically lower than for the Top 10 and 11-40 groups, at less than 30%. Moreover, the overall importance of dealers[11] in these groups diminishes as well. Specifically, in the three most-central groups, dealers (in the role of either liquidity provider or demander) account for more than 90% of trades attributed to the group. In contrast, in the less-central groups (participants 116+), dealers account for less than 40% of trades.

It is worth noting that, to my knowledge, this is the first work that speaks to the reliance on interdealer trading by peripheral dealers, i.e., outside the Top 40 participants. For groups outside the two most central groups—and specifically, the 41-115, 116-250, and 251-450 groups (i.e., the three most central groups outside the Top 40)—dealers tend to demand liquidity more often than they provide liquidity. These findings suggest that most liquidity providers (outside Top 40 dealers) tend to rely heavily on other, mainly core, liquidity providers for trading and seem to intermediate relatively little between clients. In the extreme, a peripheral liquidity provider passes on all their client trades immediately to another liquidity provider. Indeed, Skiera (2021b) shows that liquidity providers that fulfill such an intermediary role are numerous and serve many clients.

The findings of Hollifield et al. (2017), who observed all trading in the securitization market, support these observations. That study showed that peripheral dealers buy from other dealers almost 50% of the time as opposed to buying from a client. Furthermore, after buying from a client, these peripheral dealers sell to another dealer more than 50% of the time. However, that study lacked data on liquidity demand, and thus could not comment on liquidity demand. Nonetheless, both Hollifield et al. (2017) and this work provide evidence that internalizing clients' trades is only done by the most central liquidity providers, with more-peripheral liquidity providers relying heavily on other liquidity providers to reverse trades done with their clients.

[11]Dealer refers to liquidity providers. As dealers not only provide liquidity, but can also demand liquidity, the term is used to encompass both roles that dealers take and add clarity.

Summing up, the findings from Table 2.4 and Figure 2.1 show that central participants provide liquidity to less-central participants, and reveal how the core of the FX network is populated exclusively by dealers, with maybe 40 core dealers. Beyond that, few clients (non-liquidity providers) have networks that are comparable in size to the networks of large peripheral liquidity providers. Moreover, peripheral liquidity providers tend to rely heavily on core liquidity providers. They even appear more likely to pass on their trades with clients to liquidity providers, rather than holding the trades in inventory and waiting for another client to make the off-setting trade[12].

2.6 Centrality, Liquidity Provision, and Transaction Cost

As discussed in previous sections, it has long been understood that for liquidity providers to provide liquidity, they require a fee from liquidity demanders (Grossman and Miller, 1988); however, the inability to identify or infer liquidity demanders in OTC markets has prevented researchers from investigating these fees in such markets (Di Maggio et al., 2017; Hasbrouck and Levich, 2021). In this section, I study the transaction costs that a liquidity demander receives as a function of her own centrality and the centrality of the liquidity provider. Together, the results of this analysis reveal how the respective positions of a liquidity demander and her liquidity provider affect the prices that the liquidity demander trades at (conditional on requesting liquidity).

2.6.1 Transaction Cost by Centrality

Table 2.5 reports the *profit to buyer* between two centrality groups, conditional on trade initiation. I use the prefix *Buyer* when the buyer initiates the trade, by demanding liquidity (top half), and the prefix *Seller* when the seller initiates a trade. Table 2.5 immediately shows that whenever the buyer initiates a trade, the buyer pays a fee, as *profit to buyer* is always negative. Likewise, in cases where the seller demands liquidity, the seller pays a fee. Thus, the liquidity demander pays a fee to demand liquidity[13]. Unsurprisingly, the paid spread does not differ between

[12]Appendix B.2 discusses the characteristics of dealers that also demand liquidity. It shows that 97% of dealers that demand liquidity demand at least as much liquidity as they provide—leading me to conclude that I observe all liquidity demand of these dealers. Suppose that liquidity demand from other liquidity providers is not a significant part of these dealers' trading. In that case, the number of trades off the platform needs to be an order of magnitude larger than the number of trades on the platform. Given the size of the platform in the FX market, this appears unlikely.

[13]The liquidity demander does not always pay a fee; both the RFQ and Streaming summary statistics (Table 2.1) show that in more than 10% of trades the liquidity demander pays negative spreads. Negative spreads may arise when the liquidity provider has a strong imbalance in their inventory and is thus willing to off-load inventory below the midquote but (probably) within the spread. However, between any pair of participants with given centralities, the liquidity demander on average pays a fee to trade, regardless of the difference in centrality between the two parties.

buying and selling, with *profit to buyer* when a seller demands liquidity being the negative of *profit to buyer* when a buyer demands liquidity.

In Table 2.3, the buyer makes a profit if and only if the seller is less central. However, this observed pattern is purely due to the greater probability of the more-central participant providing liquidity. Therefore, the centrality premium observed in Section 2.4 simply reflects liquidity provision by more-central participants. Furthermore, controlling for liquidity provision, I observe that transaction costs vary substantially across liquidity demanders. Specifically, for a liquidity provider with a given centrality, more-central liquidity demanders receive lower transaction costs from the provider than less-central liquidity demanders receive. For example, an 11-40 participant pays around 0.2bps when demanding liquidity from an 11-40 liquidity provider, while a 41-450 participant pays around 0.3bps. And 1000+ participants pay upwards of 0.5bps to an 11-40 liquidity provider, and often above 1bps for other liquidity providers. Thus, more-central participants receive better prices when demanding liquidity.

Even more importantly, for a given liquidity demander (with the exception of the most peripheral demanders, i.e., those outside the top 450), there appears to be no pattern in the differences in transaction costs across liquidity providers. Liquidity providers in the Top 250 all provide very similar transaction costs to liquidity demanders, and it is often a toss-up who provides the lowest cost. These similar transaction costs arise despite the most central participants (Top 10) trading more than 20 times the value of a median 116-250 participant and having 15 times as many trading partners. Thus, it appears as if the centrality of the liquidity provider is unimportant to the transaction cost the liquidity demander pays. For these central centrality groups, the observed centrality premium in that table results from less-central participants demanding liquidity more often and paying a higher fee when demanding liquidity compared with more-central liquidity demanders. In other words, if a liquidity demander from a particular centrality group (centrality group i) trades with two liquidity providers (from centrality groups j and k, respectively), her transaction costs from the two providers will be similar. Conceptually, *profit to buyer* between two centrality groups i and j is:

$$\textit{profit to buyer}_{i,j} = -p_{i,j} \times c_i + \left(1 - p_{i,j}\right) \times c_j \tag{2.2}$$

where $p_{i,j}$ is the probability of i demanding liquidity from j, when trading with j; c_i is the (positive) transaction cost of i for demanding liquidity; and c_j is the (positive) transaction cost of j for demanding liquidity.

However, there are important caveats to these results: The results correspond to actual transaction fees paid by the liquidity demander, and do not capture the full spectrum of prices quoted to that demander. As noted above, the central liquidity demander contacts multiple liquidity providers rather than just one; the mean RFQ has 10 liquidity providers providing quotes. The liquidity demander then chooses the best quote among all offers. In other words, the patterns observed above are conditional on the liquidity demander accepting a quote from a given liquidity provider, and refer to the most competitive quotes. The simulation presented in the subsequent section shows the effects of this conditioning on trades.

The results are slightly different when very peripheral participants are studied (451-1000 and 1000+). These participants have much fewer trading partners—only 1.6 on average. Thus, they receive one or very few quotes when trading. Therefore, conditioning on trading changes the distribution of provided quotes much less than for central participants. For these participants, trading with a central liquidity provider (Top 10 and 11-40) tends to provide significantly lower transaction costs than trading with peripheral liquidity providers.

Despite the caveats, I conclude that the observed centrality premium in Table 2.3 is solely the result of liquidity provision. For central liquidity demanders, transaction costs are similar across liquidity providers, while for peripheral liquidity demanders, a liquidity discount is observed.

Table 2.5 is also a test of the model in Hollifield et al. (2017). Hollifield et al. (2017) argued that more-central liquidity providers provide quicker execution, allowing them to charge higher transaction fees to a particular liquidity demander compared with peripheral liquidity providers. They predict that for a given liquidity demander, transaction costs should be higher when trading with more-central liquidity providers[14]. Table 2.5 shows that this is not the case, instead providing evidence that more-central liquidity providers trade at transaction costs that are equal to or lower than those of less-central liquidity providers.

The analysis also explains the contrasting findings in Hasbrouck and Levich (2021) and Hagströmer and Menkveld (2019). As previously, Hasbrouck and Levich (2021) do not condition on trade initiation and observe a centrality premium. Hagströmer and Menkveld (2019) find at most weak evidence of a centrality premium. However, Hagströmer and Menkveld (2019) study how 8 large dealers provide quotes to clients. The study thus conditions on these dealers providing liquidity and consequently no longer finds strong evidence of a centrality premium. The different conclusions arise because the studies measure different objects. Especially, Hasbrouck and Levich (2021) measure liquidity provision by central participants as this work shows.

2.6.2 Trading Mechanism and Centrality

Given the observance of multiple different trading mechanisms, I explore how transaction costs differ between the two alternative trading mechanisms captured in the data set: RFQ and Streaming. The two trading mechanisms differ by the exclusivity in trading and revelation of

[14]The reliance on a few core liquidity providers in all OTC markets creates a common concern among regulators and market participants that the core liquidity providers use their importance in the market to force trading at high transaction costs. The results of this study suggest that this concern is not warranted in liquid markets, as the transaction costs provided by central liquidity providers are no higher than those provided by peripheral liquidity providers. However, Section 2.5 shows that peripheral liquidity providers demand a great deal of liquidity from central liquidity providers. Further studies analyzing interdealer trading are necessary to address how central liquidity providers affect the transaction costs for liquidity provision by peripheral dealers.

trading intent by the liquidity demander. In RFQ, the liquidity demander commits to executing the order with exactly one liquidity provider. Streaming, in turn, allows the liquidity demander to split the order into multiple child orders that execute simultaneously with multiple liquidity providers. However, as Panel B of Table 2.1 shows, only a few orders are split, with there being 1.074 trades per executed order and splitting occurring only at large order sizes.

Furthermore, in Streaming, the liquidity providers provide up to "last look" executable quotes to the liquidity demander, regardless of whether the liquidity demander wants to trade or not. Liquidity providers are only informed of a trade when the liquidity demander trades against their quotes, but not if the liquidity demander trades with another liquidity provider. In RFQ, the liquidity demander requests quotes from multiple liquidity providers, revealing the order size that the liquidity demander wants to trade. Subsection 2.3.2 describes each trading mechanism.

Table 2.6 shows the profit the buyer makes in a trade (*profit to buyer*), conditional on the liquidity demander, for RFQ trades. The results very much resemble the results in Table 2.5, with the liquidity demander paying a transaction cost to trade. As in Table 2.5, transaction costs are higher for more-peripheral liquidity demanders, and, for a particular demander, there is little systematic variation in the transaction costs across liquidity providers. However, the transaction costs in Table 2.6, which only covers RFQ trades, are generally higher than those in Table 2.5. For liquidity demanders in the 11-40 group, the transaction costs in RFQ trades are around 0.05bps higher than the overall costs; demanders in the 41-450 group face costs that are 0.2bps higher; and those in the 451-1000 group have costs that are 0.3bps higher. For participants in the 1000+ group, transaction costs under RFQ are more than 1bps higher than the overall costs for some centrality pairs.

Table 2.7 reports the results for trades executed via Streaming. Table 2.7, in turn, shows that while the liquidity demander still pays transaction costs to trade, these fees are lower than those in Table 2.5. This pattern is expected, given that Table 2.5 averages across trades in RFQ and Streaming. Surprisingly, Table 2.7 also shows that, in Streaming, the centrality of the liquidity demander does not matter for the transaction cost. Instead, all liquidity demanders generally receive very similar transaction costs of 0.2-0.3bps from the two most central groups of liquidity providers (Top 10 and 11-40).

Various factors may contribute to the different results under RFQ and Streaming. One possible factor is differences in the currency pairs traded through the two mechanisms. Specifically, the currency pairs traded via Streaming are a subset of the currency pairs traded via RFQ, with the least liquid currency pairs being traded only via RFQ. These currency pairs also have the highest transaction costs. However, this factor is unlikely to play a substantial role in the differences observed, given that, of the 184 currency pairs in the data, 112 currency pairs also trade via Streaming, and the low liquidity of the remaining currency pairs also means that they trade rarely. Indeed, results in Appendix B.3 show that the differences in transaction costs between the two mechanisms persist when studying only EURUSD.

A second factor that might explain the differences between the trading mechanisms is the composition of participants that tend to use each mechanism. Though Streaming makes up most of the trading volume, only 8.9% of clients (180 out of 2,024) trade via Streaming, with the remaining 1844 trading exclusively via RFQ. Most clients trade almost exclusively in either Streaming or RFQ, making a direct comparison with liquidity demander fixed effects difficult. The limited evidence from the few clients that use both trading mechanisms is that the transaction cost does not differ across these clients. Most participants that trade via Streaming have high measures of centrality. Discussions with the data provider reveal that these participants are also more sophisticated[15]. Higher sophistication allows liquidity demanders in Streaming to better evaluate whether a price is good or bad, leading to lower transaction costs; likewise, liquidity providers may be willing to provide better prices in order to trade with more-sophisticated liquidity demanders. Thus, the composition of participants may explain some of the variations in transaction cost between the trading mechanisms.

A third factor that might explain the different results obtained for the two mechanisms is differences in the trading mechanisms themselves. The two mechanisms differ in terms of the exclusivity of trading. The lack of exclusivity of trading in Streaming may cause large orders to realize lower transaction costs than they do in RFQ. Specifically, in Streaming, large orders may split across multiple liquidity providers in lower quantities and at a lower cost than they would require under trading with only one liquidity provider. However, this mechanism might be expected to lead to increased cost at smaller order sizes, as liquidity providers may find it harder to hedge an order if multiple other liquidity providers are also trying to hedge the order—forcing the provider to increase transaction costs. This situation does not arise in RFQ, where only one liquidity provider needs to hedge each trade. Since most orders are relatively small in Streaming and thus do not split into multiple orders, the lack of exclusive trading should lead to higher transaction costs in Streaming relative to RFQ.

As mentioned, a limited set of participants trade in both RFQ and Streaming, and unreported analysis indicates that, for these participants, transaction costs do not differ between the two mechanisms. This finding suggests that the differences in observed transaction costs between the trading mechanisms, and the lack of increased transaction cost for more-peripheral liquidity demanders in Streaming, are due to differences in the composition of participants in RFQ and Streaming, with participants in Streaming being more sophisticated than participants in RFQ.

2.7 Simulation: Conditioning on Trades

The previous section showed that, for all but the most peripheral liquidity demanders, transaction costs do not differ significantly across liquidity providers. I attributed these results to the demanders' simultaneous contact with many dealers and the shift in distribution when condi-

[15]Streaming is fast, with a liquidity demander getting almost immediate execution, while in RFQ, the liquidity demander first gets to observe the exact trading prices offered by the liquidity providers before deciding whether and with whom to trade.

tioning on trading. In this section, I perform a simulation to quantify these effects and to show how conditioning on trading affects the transaction cost when dealers provide different spreads. Furthermore, the simulation tries to report how differences in average spread across liquidity providers can be detected even if their effect on transaction cost across liquidity providers is small.

More formally, the simulation aims to show how differences in the spread the liquidity provider charges $s_{i,c}$ affect the transaction cost the liquidity demander realizes and the market shares that liquidity providers realize. I assume that the liquidity demander wants to buy, and that each liquidity provider i provides a quote based on the midquote $m_{c,t}$, a dealer-specific spread $s_{i,c} > 0$, and idiosyncratic shock $\epsilon_{i,c,t} \sim \mathcal{N}(0, \sigma)$, arising, for example, from inventory held by the dealer, so that a dealer's quote $q_{i,c,t}$ is given by:

$$q_{i,c,t} = m_{c,t} + s_{i,c} + \epsilon_{i,c,t} \tag{2.3}$$

Upon receiving quotes from liquidity providers $i_1, \cdots, i_N$, the liquidity demander chooses the lowest quote, i.e., $i^\star = \arg\min_{j\in\{1,\cdots,N\}} s_{i_j,c} + \epsilon_{i_j,c,t}$, giving a transaction cost of:

$$\min_{j\in\{1,\cdots,N\}} s_{i_j,c} + \epsilon_{i_j,c,t} \tag{2.4}$$

Given that EURUSD trades the most frequently (31.3% of all trades), this section is based on data using only EURUSD trades. Appendix B.3 gives a detailed description of the EURUSD market and reproduces the results of the previous sections for EURUSD trades. To represent an average transaction, the number of dealers matches the mean number of dealers in an RFQ at 10. Furthermore, the average quote by a liquidity provider has a spread of 0.32bps, which is taken as $s_{i,c}$. All that is left is to choose σ, the volatility of the quote by each dealer. σ is calibrated so that given 10 dealers, the transaction cost matches the average transaction cost in the data of 0.1bps at 10 dealers[16], i.e., $\sigma^\star$ solves:

$$0.1 = 0.32 + \mathbb{E}\left(\min_{i\in\{1,\cdots\}} \epsilon_{i,c,t}\right) \tag{2.5}$$

where $\epsilon_{i,c,t} \sim \mathcal{N}\left(0, \sigma^\star\right)$ and each i is independent from the others. This yields a value of $\sigma^\star = 0.14$[17].

Figure 2.2 shows the results of the calibration, and how improving the spreads offered by a liquidity provider affects the transaction cost a liquidity demander realizes. Panel 2.2A of Figure 2.2 shows how decreasing the spread offered by liquidity providers by 0.05bps reduces

[16]Both the average transaction cost and average quote are robust to wider bands of taking the average. For example, including all RFQ and quotes in RFQ with 6 to 14 dealers gives a mean transaction cost of 0.11 and an average quote of 0.36.

[17]The value of the expectation in equation 2.5 is linear in σ, as ϵ can be written as $\epsilon = \sigma \times X$ with X being standard normally distributed. The integral for a standard, normally distributed X is -1.54, the value in the denominator. Thus, $\sigma^\star = \frac{-0.22}{-1.54} = 0.14$

the observed transaction cost. It shows how transaction costs change for liquidity providers that reduce their spread and for liquidity providers that do not reduce their spread. As the number of liquidity providers that reduce their spread increases, the transaction cost to the liquidity demander decreases. However, Panel 2.2A also shows that when a demander trades, transaction costs with both types of liquidity providers decline. This reduction in transaction cost arises from the provided quotes being of lower cost.

Furthermore, when dealers reduce their spread by 0.05bps, the overall transaction cost, i.e., conditioning on trading with a liquidity provider, is only 0.013-0.014bps lower compared to that obtained with liquidity providers that did not lower their spread. This low pass-through of reduction of the spread to the overall transaction cost arises from the fact that the transaction cost only reflects the best quote, which is the quote that leads to a trade. Conditioning on providing the best quote among ten liquidity providers will produce very similar transaction costs distributions across heterogeneous liquidity providers, as 9 of the 10 liquidity providers will be the same when comparing two different liquidity providers. Thus, quotes need to cross a similar threshold across these liquidity providers. These results illustrate how simply focusing on observed trades underestimates any differences in the spreads that liquidity providers charge liquidity demanders.

Panel 2.2B of Figure 2.2 shows, however, a margin where liquidity providers are different, namely their probability of trading with a liquidity demander. While the average spreads at which liquidity providers trade differ from one another by only 14%($= \frac{0.014}{0.1}$), the probabilities of trading with a given liquidity demander differ by more than 70%($= \frac{0.165-0.093}{0.1}$). A liquidity provider who offers a better spread than others is likely to provide the best quote much more often than other liquidity providers—even in cases where previously her inventory shock $\epsilon_{i,c,t}$ would have meant that she would not have provided the best quote. Panels 2.2C and 2.2D also show these results. Panels 2.2C and 2.2D study the effects of having each dealer provide a 0.01bps lower spread than the next dealer. (The average spread offered is shown on the x-axis.) Even when comparing the two most different dealers, the difference in transaction cost is again very small, 0.026bps, compared to the difference in average spread (0.09bps). However, the difference in probability in trading remains large, with the liquidity provider that provides the lowest average spread trading three times more often than the liquidity provider with the largest average spread.

The simulations show that it is not surprising that no pattern is found across liquidity providers once I condition on realizing a trade. Given that many quotes are provided simultaneously to a liquidity demander, conditioning on trades implies that only cases in which a liquidity provider provided the best quote are measured. In these cases, any differences in the average spread provided by dealers are reduced dramatically. Differences as a function of centrality are thus more likely to be observed in the probability of trading with a given liquidity provider or the share of trading done with a specific liquidity provider rather than in the trade prices at which a liquidity demander trades with a liquidity provider.

The simulation and empirical results in Tables 2.4 and 2.5 reveal how constraints of data availability might limit the insights that can be obtained regarding relationships between network structure and trading outcomes. First, the empirical results show how the centrality premium observed in previous studies (as well as in section 2.4) captures mainly liquidity provision from more-central participants to less-central participants. This finding highlights the need to be able to identify liquidity providers versus demanders in OTC markets, particularly since it is well understood that demanding liquidity requires the liquidity demander to pay a fee. Yet, the results of the simulation show that in liquid markets, like the FX market, even when information on liquidity provision is available, the fact that observation of prices is conditioned on trades further limits any insights that can be obtained[18]. This limitation arises from the fact that conditioning on the best available price forces any differences in liquidity providers' average spread to be very much compressed.

The following section proposes an alternative approach to evaluating the relationship between costs and trading costs—namely, focusing on the share of trading with each liquidity provider rather than solely on costs. This focus is motivated by the observation that, as shown in the simulation, differences in average spread have a more significant effect on the share of trading done with the respective liquidity provider compared to the transaction cost.

2.8 Centrality, Liquidity Provision, and Market Share

Table 2.5 shows that peripheral liquidity demanders enjoy a centrality discount. This finding hints that a centrality discount might exist more broadly for all liquidity demanders. However, because central liquidity demanders receive many quotes simultaneously, looking only at trades implies that only the best quotes are observed. Therefore, a centrality discount (in terms of the spreads offered by liquidity providers) may not translate to transaction costs. The simulation above verifies these intuitions and proposes that looking at market shares would give a clearer picture. The intuition is that the liquidity provider supplying the best quotes on average will trade most often, though the provider may not necessarily trade at lower transaction costs compared with other providers. In other words, when I focus on market shares, I can conclude that a centrality discount exists when more-central liquidity providers trade relatively more frequently than less-central liquidity providers, even if the transaction costs are similar.

For each liquidity demander, I determine the market share $f_{j,i}$ of each liquidity provider i the liquidity demander j trades with as follows:

$$f_{j,i} = \frac{\#Trades\ between\ j\ and\ i\ where\ j\ demands\ liquidity\ from\ i}{\#Trades\ where\ j\ demands\ liquidity} \tag{2.6}$$

[18]Even if all requested quotes are provided and observed, liquidity demanders may only contact a liquidity provider is a subset of currency pairs, as other provided quotes may not be competitive enough. In that case, the average spread would still be biased as the liquidity provider is only observed when the spread is competitive. If liquidity providers selectively provide quotes, the observed average spread may again be biased.

Central liquidity providers have the most liquidity demanders. However, that does not imply that a liquidity demander trades more with a central than a peripheral liquidity provider. I relate the market share of the liquidity provider to the centrality of the liquidity provider, as follows:

$$\log\left(f_{j,i}\right) = \beta \log\left(\#Liquidity\ Provider_j\right) + \gamma_{Centrality_i} + \epsilon_{j,i} \tag{2.7}$$

where $\log\left(\#Liquidity\ Provider_j\right)$ is the log number of liquidity providers that liquidity demander j trades with, and $\gamma_{Centrality_i}$ are a set of dummy variables for each centrality group. The dummy corresponding to the centrality group in which liquidity provider i is a member is going to be equal to 1. One observation is one trading relationship between a liquidity demander and a liquidity provider.

I include a liquidity demander's log-number of liquidity providers, as a larger number of liquidity providers means a lower average market share for each liquidity provider (average market share $\left(= \frac{1}{\#Liquidity\ Provider}\right)$. Thus, $-\log\left(\#Liquidity\ Provider_j\right)$ represents the log of the average market share.

Table 2.8 shows the regression results in specification (1). The coefficient on the log number of liquidity providers is -1.25. The logarithm is concave, so that

$$\mathbb{E}\left(\log\left(f_{j,i}\right)\right) \leq \log\left(\mathbb{E}\left(f_{j,i}\right)\right) = -\log\left(\#Liquidity\ Provider_j\right) \tag{2.8}$$

Therefore, the coefficient on $\log\left(\#Liquidity\ Provider_j\right)$ is at most -1 and is only equal to -1 if all liquidity providers provide the same amount of liquidity to a liquidity demander. The difference between the coefficient and -1 measures the dispersion of market shares. The coefficients on the liquidity provider's centrality group show that more-central liquidity providers trade relatively more with a liquidity demander than do peripheral liquidity providers. A Top 10 liquidity provider trades 37.7%($=e^{0.32}-1$) more with a liquidity demander than does the liquidity demander's average liquidity provider, where an average liquidity provider has the average log-market share of liquidity demander's liquidity providers. Liquidity providers in the 11-40 group trade proportionally to the average log-market share, while peripheral liquidity providers trade less than the average liquidity provider. For example, 41-115 liquidity providers trade 15.7%($=\left|e^{-0.17}-1\right|$) less than the average liquidity provider, with 1000+ liquidity providers trading 28.8%($=\left|e^{-0.34}-1\right|$) less than the average liquidity provider. These results show that a liquidity demander trades much more with her central liquidity providers than with her peripheral liquidity providers. The results support a centrality discount in the sense that a liquidity demander receives a better average price from more-central liquidity providers, leading them to trade more often with their more-central liquidity providers.

Table 2.9 repeats the analysis of Table 2.8, except for reporting liquidity demander by liquidity provider fixed effects. Therefore, the table shows how the market share corresponding to liquidity providers in a given centrality group varies across liquidity demanders. Rows represent the centrality of the liquidity demander, and columns represent the centrality of the liquidity provider. The coefficient on the log number of liquidity providers is now -1.04 (see Table description), statistically no different from -1. This result implies that a liquidity demander demands similar liquidity from all liquidity providers in a centrality group. Differences thus exist in how much trading providers in different centrality groups execute with a particular liquidity demander relative to the average liquidity provision. Figure 2.1 shows that participants in the 11-40 group demand very little liquidity. Therefore, many coefficients are imprecisely estimated for this group. However, for liquidity demanders in the 41-1000 group, I find that more-central liquidity providers have a higher market share compared with peripheral liquidity demanders. For example, for a 41-115 liquidity demander, the coefficient for a Top 10 liquidity provider is 0, whereas the coefficient for a 251-450 provider is -1.6, meaning that a 251-450 liquidity provider provides 80%(=$e^{-0.0-1.6} - 1$) less liquidity than a Top 10 liquidity provider.

The table also shows that the differences in provided liquidity decrease the more peripheral the liquidity demander is. When a liquidity demander receives prices from 10 liquidity providers, the liquidity demander likely tolerates trading with one of the liquidity providers only 5% of the time (50% below the average market share). However, if a liquidity demander has only 4 liquidity providers, trading 12.5% of the time (again 50% below the average market share) is likely too costly to maintain the relationship. Indeed, unreported results show that the standard deviation in liquidity providers' log market shares is increasing in the centrality of the liquidity demander.

For the most peripheral liquidity demanders, I find a centrality discount when looking at transaction cost (Table 2.5). At the same time, there is no difference in market share based on the centrality of the liquidity provider, which is likely to be attributable to the fact that many of these liquidity demanders only receive liquidity from one liquidity provider.

Summing up, Table 2.5 shows that a centrality discount exists for peripheral liquidity demanders, as conditioning on trading does not affect which quotes from a liquidity provider are observed. Table 2.9 shows that for central liquidity demanders, conditioning on trading leads to observing many more trades with central liquidity providers than with peripheral liquidity providers. This conditioning significantly changes the observed quotes, leading to a lack of a centrality discount when studying prices (Table 2.5). However, taken together, Table 2.5 and Table 2.8 and 2.9 support the proposition that a centrality discount exists, i.e., more central liquidity providers provide better average prices to liquidity demanders.

2.9 Conclusion

OTC markets are organized in networks with few core participants and many peripheral participants. Understanding how a participant's network positioning affects the pricing the participant receives is key to understanding the price dispersion in and functioning of OTC markets. Previous studies described OTC networks in great detail, in numerous types of OTC markets (Hollifield et al., 2017; Di Maggio et al., 2017; Li and Schürhoff, 2019; Hasbrouck and Levich, 2021), and measure how positioning affects pricing. However, these studies did not observe a key variable of major importance for pricing, namely, trade initiation. My analysis overcame this limitation by working with data from a leading multi-dealer platform in the FX market, in which each trade identified the liquidity provider and the liquidity demander, thereby enabling me to observe trade initiation.

I first analyzed the data without taking into account the extra information on trade initiation, and showed that the data replicated the findings of Hasbrouck and Levich (2021)—revealing a centrality premium, in which the more-central participant in a trade tends to make a profit.

Next, utilizing my information on trade initiation, I showed that more-central participants generally provide liquidity to less-central participants. These dynamics explain the centrality premium observed in the previous step (and in previous studies)—that is, the centrality premium captures liquidity provision by central participants. Liquidity demand, in turn, requires all participants, regardless of centrality (their own or their trading counterparts'), to pay a transaction cost. However, central liquidity demanders enjoy lower transaction costs compared with peripheral liquidity demanders.

Next, controlling for liquidity provision, I examined the relationship between a liquidity provider's centrality and the transaction fees she charges. In this analysis, no centrality premium emerged. Rather, for peripheral liquidity demanders, I observed a centrality discount, in which more-central liquidity providers provide lower transaction costs. For other, more-central, liquidity demanders, peripheral and central liquidity providers provide similar prices.

I explained the differences between less- and more-central liquidity demanders on the basis of the fact that my analysis was conditioned on the realization of a trade. In other words, the observed prices were the most competitive quotes that liquidity demanders were able to obtain, rather than the full spectrum of quotes they were offered. Compared with peripheral liquidity demanders, more-central liquidity demanders are able to contact many more liquidity providers, and thus operate in a more competitive environment—potentially reducing the variation in prices at which they ultimately trade. Indeed, a simulation confirmed this proposition, showing that differences across liquidity providers in terms of quotes provided translate into much smaller differences in terms of prices realized. For example, the results of the simulation suggest that if a liquidity provider provides better prices, on average, compared with other liquidity providers, the provider's advantage is reduced by 70% when only trades are observed.

However, the simulation also shows that better average prices have a much greater effect on the market share of a liquidity provider than on the transaction fees she charges. For example, I find that a Top 10 liquidity provider trades more than 90%($=e^{0.32-(-0.34)}-1$) as much with a liquidity demander than a peripheral (1000+) liquidity provider, even if the liquidity demander requests liquidity from both. More generally, the more central the liquidity provider, the larger the market share of the liquidity provider, even after controlling for the average market share of the liquidity demander's liquidity providers. Thus, I find that a centrality discount exists, in that more-central liquidity providers provide better prices on average.

This study shows that the FX market exhibits a centrality premium when one does not assign trade initiation. However, to my knowledge, this study is the first also to provide concrete evidence for why the centrality premium arises: Namely, because more-central participants tend to provide liquidity, whereas less-central participants initiate trades. Once I control for trade initiation, a centrality discount emerges with more-central liquidity providers supplying better prices on average.

In practical terms, these findings may be useful to clients trading in OTC markets. These clients are not in a position to offer quotes to other participants and thus only demand liquidity. Their decision power is limited to determining how many and which dealers to trade with. In revealing the relationship between a liquidity provider's centrality and the transaction fees that the provider is likely to charge, this study may assist clients in constructing their trading networks wisely.

2.10 Chapter 2: Figures

Figure 2.1: Share of Liquidity Provision by Centrality Group: This figure shows the share of trades where a centrality group demands/provides liquidity. Navy shows the share of trades in which the dealers in a given centrality group provide liquidity. Dark green shows the share of trades in which these dealers demand liquidity. Light green shows the share of trades by non-liquidity providers; these are all trades where the non-liquidity providers demand liquidity. If a single Streaming order splits into multiple trades, each trade counts as a single observation in the figure. The word "dealers" refers to "liquidity providers".

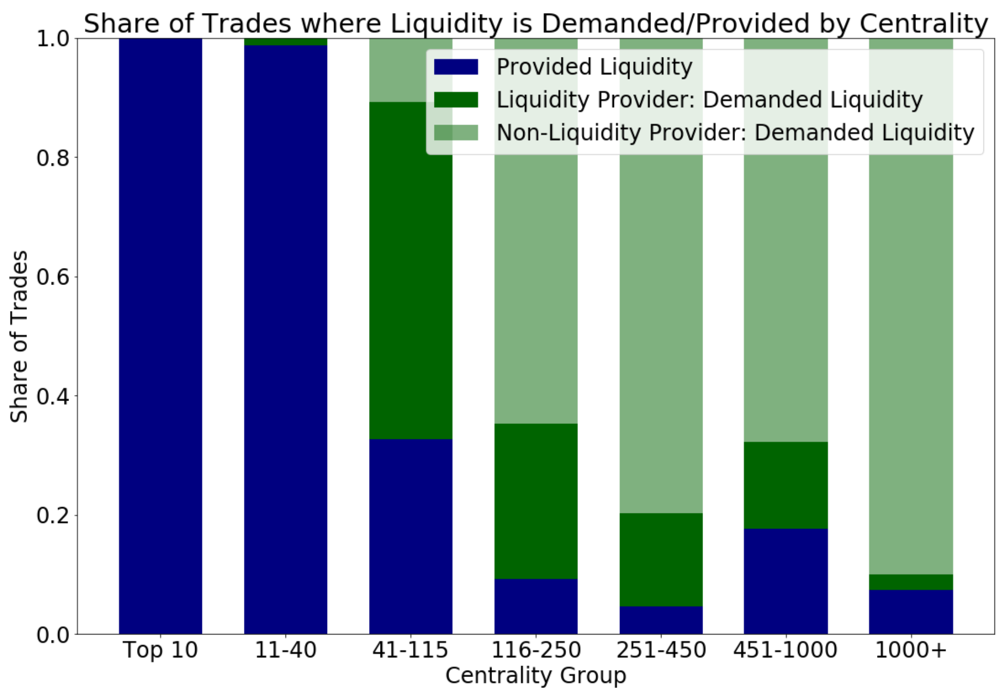
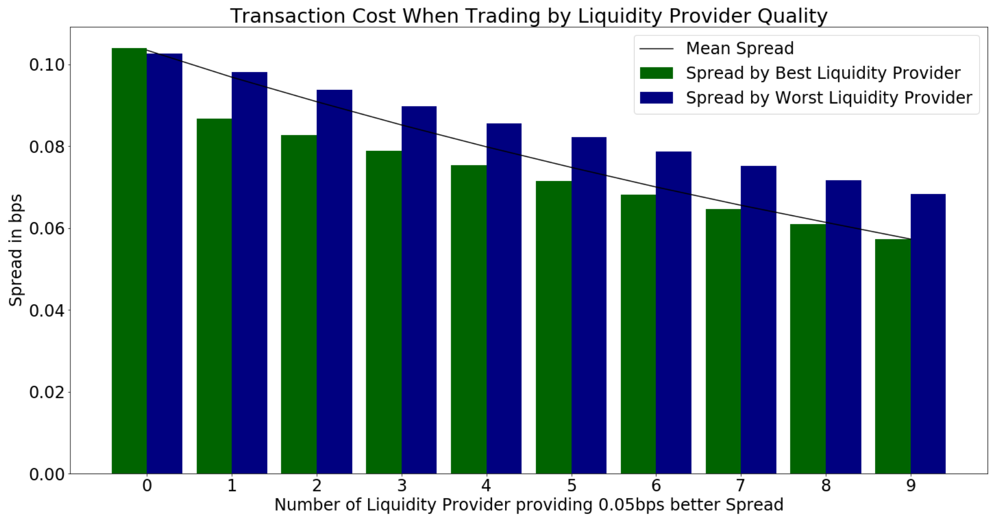
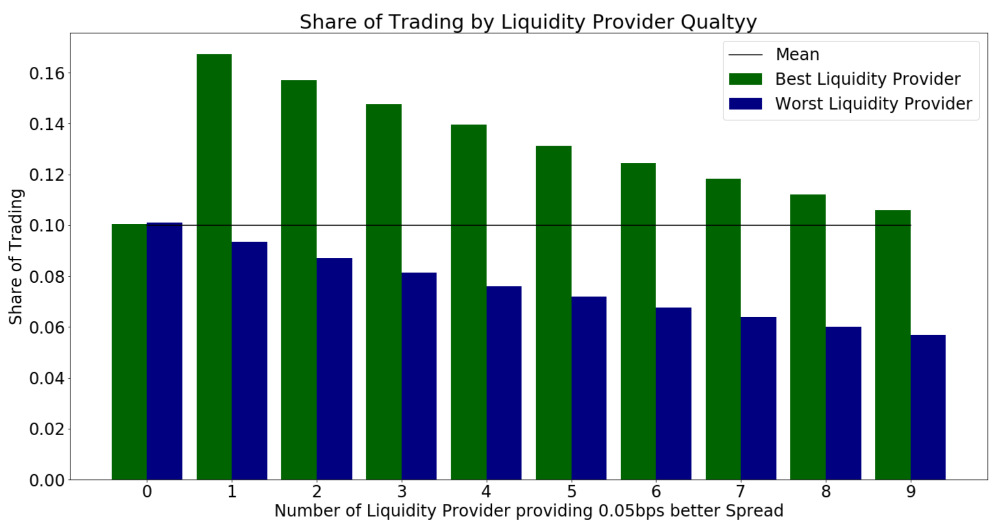


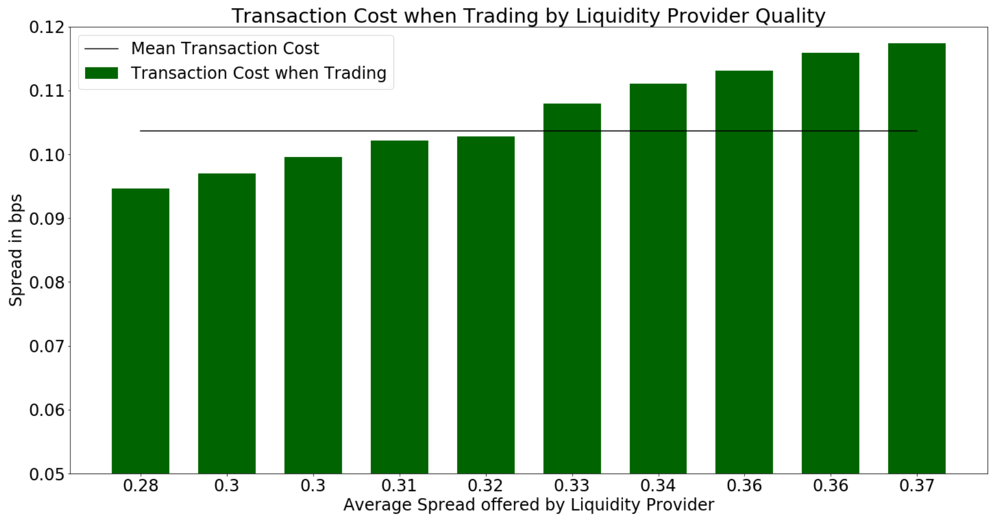
Figure 2.2: Simulation of Transaction Cost and Share of Trading with Heterogeneous Liquidity Providers: This table shows the simulated transaction cost and share of trading for heterogeneous liquidity providers. Quotes are determined to be offered according to $q_{i,c,t} = m_{c,t} + s_{i,c} + \epsilon_{i,c,t}$, where $s_{i,c}$ is the average spread offered by the liquidity provider (calibrated to 0.32 in the baseline) and $\epsilon \sim \mathcal{N}(0, 0.14)$. In each simulation there are 10 dealers (the average number in an RFQ). **Panel** 2.2A **and** 2.2B: show the resulting transaction cost and share of trading (respectively) as the number of liquidity providers offering lower spreads increases. Liquidity providers offering a lower spread offer a spread of 0.27bps on average, 0.05bps lower than the liquidity providers offering the high spread. The number of liquidity providers offering the low spread increases from 0 to 9 and reports the outcomes for liquidity providers offering the lower spread (navy, *Low Spread Liquidity Provider*) and liquidity provider offering the higher spread (0.32bps) (dark green, *High Spread Liquidity Provider*. **Panel** 2.2C **and** 2.2D show the transaction cost and share of trading (respectively) for heterogeneous liquidity providers, offering average spreads of 0.28bps up to 0.37bps. Each subsequent liquidity provider offers a spread 0.01bps higher than the previous liquidity provider. Reported are the transaction cost and share of trading for each type of dealer.



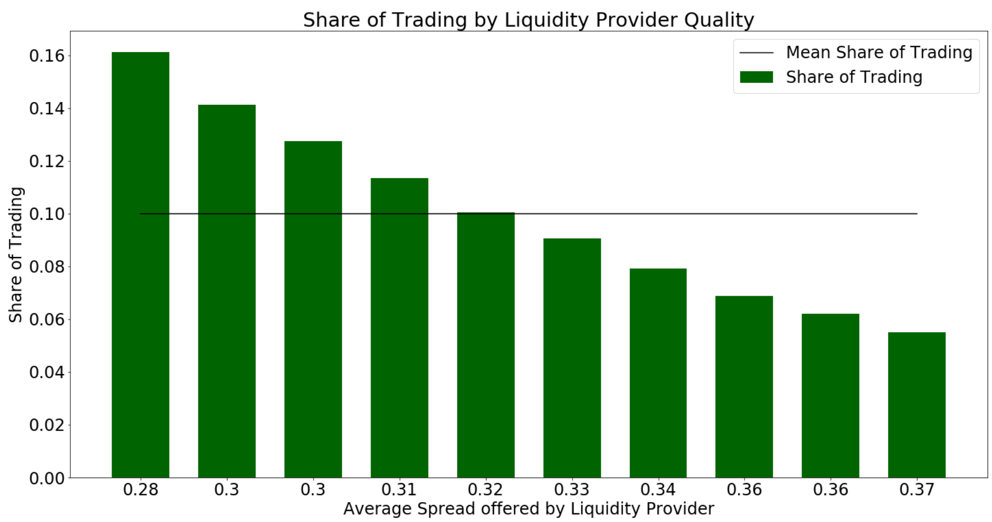
Panel A: Transaction Cost



Panel B: Share of Trading



Panel C: Transaction Cost



Panel D: Share of Trading

2.11 Chapter 2: Tables

Table 2.1: Summary Statistics: This table shows summary statistics for order and trade sizes, as well as transaction costs, in RFQ and Streaming. **Panel A** reports summary statistics for RFQ, and **Panel B** reports summary statistics for Streaming. **Panel A**: *RFQ: Order Size* reports the distribution of order size for all RFQ, while *Trade Size* reports only the distribution of order size for RFQ that end in a trade. Note that in each RFQ, only one quote gets executed. *Trades: Half-Spread* reports the distribution of the spread, between the midquote and the trade price the liquidity demander trades at, measured in log bps. **Panel B**: *Orders: Order Size* reports the distribution of order size across all orders in Streaming, while *Executed Orders: Order Size* restricts the distribution to those orders that have at least one execution. *Executed Orders: Filled Size* reports the distribution of the amount traded across trades affiliated with an order that has at least one execution. *Trade Size* reports the distribution of amount traded per trade an executed order has. *Trades: Half-Spread* reports the distribution of the spread, between the midquote and the trade price the liquidity demander trades at, measured in log bps.

Panel A: Summary Statistics Request for Quote (RFQ)

	N	Mean	Std.	10%	25%	50%	75%	90%	95%
RFQ: Order Size	637,816	1.45	2.75	0.13	0.22	0.50	1.25	3.61	5.59
Trade Size	404,065	1.50	2.98	0.13	0.22	0.50	1.22	3.50	6.35
Trades: Half-Spread	398,800	0.75	1.40	-0.04	0.09	0.29	0.89	1.88	2.90

Panel B: Summary Statistics Streaming

	N	Mean	Std.	10%	25%	50%	75%	90%	95%
Orders: Order Size	962,041	1.05	1.68	0.18	0.34	0.64	1.00	1.85	4.40
Executed Orders: Order Size	927,602	1.06	1.7	0.18	0.34	0.67	1.00	1.98	4.43
Executed Orders: Filled Size	927,602	1.05	1.67	0.18	0.34	0.67	1.00	1.85	4.00
Trade Size	996,531	0.99	1.50	0.19	0.41	0.89	1.00	1.79	3.00
Trades: Half-Spread	908,237	0.31	0.74	-0.09	0.02	0.16	0.35	0.73	1.17

Table 2.2: Descriptive Statistics of Participants: This table shows the distribution of degree centrality and trade volume by participant centrality. A participant's centrality is defined as her degree centrality, i.e., the number of trading partners she has. **Panel A:** shows the distribution of degree centrality both across all participants and by each group of centrality. **Panel B:** shows the distribution of trade volume across all participants and within each group of centrality. Trade volume is in mil EUR. I form the groups of centrality so that the four most central groups (measured by high degree centrality) cover around 40% of trading each, the fifth group covers around 20% of trading, and the final two groups cover around 10% of trading each.

Panel A: Distribution of Degree Centrality across Participants

	N	Mean	Std.	10%	25%	50%	75%	90%	Max
All	2,411	8.7	27.6	1.0	1.0	3.0	7.0	16.0	449.0
Top 10	10	330.4	66.1	259.9	280.5	312.0	370.8	416.6	449.0
11-40	30	137.6	48.7	74.0	97.5	132.5	167.0	215.3	234.0
41-115	75	41.2	11.8	29.0	31.0	39.0	48.0	61.2	68.0
116-250	135	19.6	3.3	16.0	17.0	19.0	22.0	25.0	27.0
251-450	200	11.7	1.8	9.0	10.0	12.0	13.0	14.0	15.0
451-1000	550	5.8	1.5	4.0	4.0	6.0	7.0	8.0	9.0
1000+	1,411	1.5	0.8	1.0	1.0	1.0	2.0	3.0	4.0

Panel B: Distribution of Trade Volume across Participant's Centrality

	N	Mean	Std.	10%	25%	50%	75%	90%	Max
All	2,411	1,333.5	6,301.4	1.0	5.6	32.9	223.2	1,958.9	117,831.5
Top 10	10	67,481.7	25,652.2	47,319.9	48,745.2	54,525.5	86,193.8	98,877.6	117,831.5
11-40	30	20,822.0	19,284.1	5,730.3	7,090.1	15,189.3	26,812.8	40,967.6	79,339.6
41-115	75	8,686.1	10,540.2	1,099.5	2,323.6	4,628.2	10,824.5	21,461.4	51,920.0
116-250	135	4,600.0	5,858.0	347.0	713.5	1,801.5	6,597.8	13,043.3	28,336.2
251-450	200	1,605.5	3,137.4	93.5	178.4	403.9	1,462.2	3,672.0	19,581.4
451-1000	550	331.3	935.2	19.1	35.0	81.8	222.7	637.6	10,604.5
1000+	1,411	99.1	511.1	0.5	1.9	8.5	30.8	123.7	9,653.5

Table 2.3: Profit to Buyer by Centrality: This table looks at the profit that a buyer with a certain centrality (row) makes when trading with a seller with a certain centrality (column). *Profit to buyer* is measured as the difference in log basis points between the midquote for the currency pair and the price the buyer pays to buy the quote currency for the base currency, i.e., the half-spread paid by the buyer. (See Equation 2.1 for the definition.) The reported values are the regression coefficient where the profit of a certain buyer is regressed on *buyer centrality* × *seller centrality* dummy variables, as well as *Date*, *Time of Day*, *Trade mechanism* (RFQ or Streaming) and *Currency Pair* fixed effects. For ease of interpretation, only *buyer centrality* × *seller centrality* dummy variables are reported in table form, with the row representing the centrality of the buyer and the column representing the centrality of the seller. The regression includes 1,310,865 observations, constructed by all RFQ and Streaming trades with an order size of at least 100,000 EUR. (See Appendix B.1 for a detailed description of the construction of the sample.) The standard errors are reported below the regression coefficient. Standard errors are clustered at the *Currency Pair* × *Time of Date* level. *Time of Date* is a 30-minute interval on a particular date.

Profit to Buyer by Centrality

Buyer\Seller	Top 10	11-40	41-115	116-250	251-450	451-1000	1000+
Top 10		0.17***	0.28***	0.25***	0.27***	0.37***	0.68***
		(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
11-40	-0.12***	0.06	0.31***	0.30***	0.39***	0.43***	0.53***
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
41-115	-0.30***	-0.34***	0.00	0.15***	0.27***	0.50***	1.17***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
116-250	-0.25***	-0.27***	-0.16***	-0.01	0.18***	-0.1***	1.35***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
251-450	-0.26***	-0.38***	-0.23***	-0.11***	0.16***	0.82***	0.70***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)
451-1000	-0.35***	-0.45***	-0.50***	-0.09**	-0.46***	-0.01	2.75***
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)
1000+	-0.73***	-0.54***	-1.00***	-0.86***	-0.93***	-0.74***	0.13**
	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)

Table 2.4: Liquidity Demand between Centrality Pairs: This symmetric table shows the share of trades between two participants with centrality corresponding to a given *row* and *column*, in which the *row*-centrality participant demands liquidity from the *column*-centrality participant.

Liquidity Demand between Centrality Pairs							
	Top 10	11-40	41-115	116-250	251-450	451-1000	1000+
Top 10							
11-40	1.0	0.5					
41-115	1.0	1.0	0.5				
116-250	1.0	1.0	0.86	0.5			
251-450	1.0	1.0	0.91	0.67	0.5		
451-1000	1.0	0.98	0.84	0.31	0.63	0.5	
1000+	1.0	1.0	0.99	0.99	0.68	0.87	0.5

Table 2.5: Profit to Buyer by Centrality Conditional on Liquidity Demand: This table looks at the profit a buyer with a certain centrality makes when trading with a seller with a certain centrality. The rows indicate the liquidity-requesting participant, i.e., the trade initiator. In cases with the prefix *Buyer*, the row represents the buyer in the transaction, and the column represents the seller in the transaction. In cases with the prefix *Seller*, these are switched, with the row representing the seller and the column representing the buyer. However, unchanged is that each cell represents the profit to the buyer. Profit is measured as the difference in log basis points between the price at which the buyer buys the currency pair and the midquote for the currency pair, i.e., the half-spread paid by the buyer. The reported values are the regression coefficient where I regress the profit of a certain buyer on buyer centrality x seller centrality dummy variables, conditional on the initiation of the trade. The regression includes 1,310,865 observations and further includes Date and Time of Day fixed effects, Currency Pair fixed effects, trade mechanism fixed effects, and trade size controls (linear and log of trade size). The coefficients on trade size (in mil) are: Linear -0.0 (0.0) and Log 0.0 (0.0). The sample is constructed looking at all RFQ and Streaming trades with a trade size of at least 100,000 EUR. For Streaming trades, the trade size is the order size, not just the amount traded with the dealer in that trade. The standard errors are reported below the regression coefficient. Standard errors are clustered at the *Currency Pair* $\times$ *Time of Date* level.

Profit to Buyer by Centrality Conditional on Liquidity Demand

	Top 10	11-40	41-115	116-250	251-450	451-1000	1000+
Buyer: Top 10							
Buyer: 11-40	-0.11***	-0.20***	-0.12***	-0.13***		-0.04	
	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	
Buyer: 41-115	-0.30***	-0.33***	-0.28***	-0.42***	-0.75***	-0.39***	-0.25***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Buyer: 116-250	-0.24***	-0.27***	-0.23***	-0.24***	-0.64***	-0.52***	-0.37***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.06)
Buyer: 251-450	-0.25***	-0.37***	-0.32***	-0.38***	-0.97***	-0.39***	-0.86***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.04)
Buyer: 451-1000	-0.34***	-0.46***	-0.67***	-1.04***	-1.42***	-0.44***	-1.05***
	(0.03)	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)
Buyer: 1000+	-0.72***	-0.53***	-1.01***	-0.87***	-1.70***	-1.00***	-1.93***
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
Seller: Top 10							
Seller: 11-40	0.18***	0.23***	0.19***	0.22***		0.13***	
	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	
Seller: 41-115	0.29***	0.32***	0.29***	0.33***	0.62***	0.35***	0.34***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.06)	(0.04)
Seller: 116-250	0.26***	0.31***	0.25***	0.23***	0.51***	0.40***	0.53***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.09)
Seller: 251-450	0.28***	0.40***	0.37***	0.59***	1.10***	0.34***	0.98***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)	(0.06)
Seller: 451-1000	0.38***	0.45***	0.64***	0.91***	1.20***	0.35***	1.30***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Seller: 1000+	0.69***	0.54***	1.19***	1.38***	1.65***	3.35***	2.24***
	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)

Table 2.6: Profit to Buyer by Centrality Conditional on Liquidity Demand in RFQ Trades: This table looks at the profit that a buyer with a certain centrality makes when trading with a seller with a certain centrality. The rows are from the view of the liquidity-requesting participant, i.e., the trade initiator. In cases with the prefix *Buyer*, the row represents the buyer in the transaction and the column represents the seller of the transaction. In cases with the prefix *Seller*, these are switched, with the row representing the seller and the column representing the buyer. However, unchanged is that each sell represents the profit to the buyer. Profit is measured as the difference in log basis points between the price at which the buyer buys the currency pair and the midquote for the currency pair, i.e., the half-spread paid by the buyer. The values are the regression coefficient where I regress the profit of a certain buyer on buyer centrality x seller centrality dummy variables. The regression includes 402,628 observations and further includes Date and Time of Day fixed effects, Currency Pair fixed effects, trade mechanism fixed effects, and trade size controls (linear and log of trade size). The coefficients on trade size (in mil) are: Linear -0.01 (0.0) and Log 0.01 (0.0). The sample is constructed by looking at RFQ trades with a trade size of at least 100,000 EUR. Standard errors are clustered at the *Currency Pair* $\times$ *Time of Date* level.

Profit to Buyer by Centrality Conditional on Liquidity Demand in RFQ Trades

	Top 10	11-40	41-115	116-250	251-450	451-1000	1000+
Buyer: Top 10							
Buyer: 11-40	-0.24***	-0.23***	-0.18**	-0.21***		-0.12	
	(0.08)	(0.08)	(0.08)	(0.08)		(0.07)	
Buyer: 41-115	-0.55***	-0.5***	-0.4***	-0.61***	-0.93***	-0.35***	-0.41***
	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)
Buyer: 116-250	-0.44***	-0.46***	-0.32***	-0.35***	-0.51***	-0.25***	-0.47***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.1)
Buyer: 251-450	-0.49***	-0.57***	-0.47***	-0.43***	-1.16***	-0.48***	-0.93***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.09)	(0.08)
Buyer: 451-1000	-0.58***	-0.57***	-0.96***	-1.18***	-1.64***	-1.04***	-1.15***
	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.1)	(0.08)
Buyer: 1000+	-1.23***	-1.9***	-1.63***	-1.01***	-1.79***	-2.09***	-2.23***
	(0.08)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)
Seller: Top 10							
Seller: 11-40	0.14*	0.14*	0.12	0.16**		0.06	
	(0.07)	(0.07)	(0.08)	(0.07)		(0.07)	
Seller: 41-115	0.26***	0.23***	0.22***	0.27***	0.49***	0.2*	0.29***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.1)	(0.11)
Seller: 116-250	0.27***	0.31***	0.21***	0.17**	0.51***	0.16**	0.48***
	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.11)
Seller: 251-450	0.38***	0.47***	0.33***	0.48***	0.97***	0.24***	0.9***
	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.09)
Seller: 451-1000	0.56***	0.51***	0.85***	0.88***	1.07***	0.67***	1.23***
	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Seller: 1000+	1.01***	1.47***	1.78***	1.84***	1.57***	1.47***	2.46***
	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)

Table 2.7: Profit to Buyer by Centrality Conditional on Liquidity Demand in Streaming Trades: This table looks at the profit a buyer with a certain centrality makes when trading with a seller with a certain centrality. The rows are from the view of the liquidity-requesting participant, i.e., the trade initiator. In cases with the prefix *Buyer*, the row represents the buyer in the transaction and the column represents the seller of the transaction. In cases with the prefix *Seller*, these are switched, with the row representing the seller and the column representing the buyer. However, unchanged is that each sell represents the profit to the buyer. Profit is measured as the difference in log basis points between the price at which the buyer buys the currency pair and the midquote for the currency pair, i.e., the half-spread paid by the buyer. The reported values are the regression coefficient where I regress the profit of a certain buyer on buyer centrality x seller centrality dummy variables, conditional on the initiation of the trade. The regression includes 908,237 observations and further includes Date and Time of Day fixed effects, Currency Pair fixed effects, trade mechanism fixed effects, and trade size controls (linear and log of trade size). The coefficients on trade size (in mil) are: Linear 0.0 (0.0) and Log -0.01 (0.0). The sample is constructed looking at Streaming trades with a trade size of at least 100,000 EUR. For Streaming trades, the trade size is the order size, not just the amount traded with the dealer in that trade. Standard errors are clustered at the *Currency Pair* $\times$ *Time of Date* level.

Profit to Buyer by Centrality Conditional on Liquidity Demand in Streaming Trades

	Top 10	11-40	41-115	116-250	251-450	451-1000	1000+
Buyer: Top 10							
Buyer: 11-40	-0.07*	-0.3***	-0.29**	-0.07		-0.0	
	(0.04)	(0.07)	(0.12)	(0.2)		(0.06)	
Buyer: 41-115	-0.24***	-0.32***	-0.28***	-0.32***	-0.65***	-0.47***	-0.28***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.04)
Buyer: 116-250	-0.23***	-0.28***	-0.25***	-0.22***	-0.65***	-0.5***	-0.37***
	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.07)
Buyer: 251-450	-0.22***	-0.34***	-0.31***	-0.51***	-0.74***	-0.43***	
	(0.03)	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)	
Buyer: 451-1000	-0.26***	-0.45***	-0.43***	-0.95***	-0.49***	-0.24***	
	(0.04)	(0.04)	(0.05)	(0.07)	(0.07)	(0.04)	
Buyer: 1000+	-0.22***	-0.27***	-0.37***	-0.73***		0.05	-0.57***
	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	(0.04)
Seller: Top 10							
Seller: 11-40	0.05	0.25***	0.28	-0.03		-0.01	
	(0.05)	(0.05)	(0.19)	(0.07)		(0.06)	
Seller: 41-115	0.22***	0.3***	0.25***	0.26***	0.63***	0.45***	0.3***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.05)
Seller: 116-250	0.23***	0.26***	0.21***	0.17***	0.58***	0.47***	0.42***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.1)
Seller: 251-450	0.22***	0.32***	0.32***	0.64***	1.11***	0.31***	
	(0.04)	(0.04)	(0.04)	(0.06)	(0.09)	(0.05)	
Seller: 451-1000	0.18***	0.33***	0.24***	0.66***	0.62***	0.14***	
	(0.04)	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	
Seller: 1000+	0.19***	0.23***	0.27***	0.64***		4.76***	0.52***
	(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	(0.04)

Table 2.8: Log Market Share by Liquidity Provider: This table looks at the market share of a liquidity provider centrality group. The table reports the regression coefficients when the log market share of a liquidity provider to a liquidity demander is regressed on the liquidity demander's log number of liquidity providers and the centrality of the liquidity provider. I calculate market share as the share of trades a liquidity provider executes with a liquidity demander as a fraction of all the liquidity demander's trades. The sample of trades is constructed by looking at Streaming and RFQ trades with a trade size of at least 100,000 EUR. For Streaming trades, the trade size is the order size, not just the amount traded with the dealer in that trade.

Log Market Share by Liquidity Provider	
	(1)
Log Liquidity Provider	-1.25***
	(0.01)
Top 10	0.32***
	(0.03)
11-40	-0.02
	(0.03)
41-115	-0.17***
	(0.03)
116-250	-0.31***
	(0.06)
251-450	-0.34***
	(0.07)
451-1000	-0.05
	(0.07)
1000+	-0.34***
	(0.09)
N	10,438
R^2	0.5215

Table 2.9: Log Market Share of Liquidity Provider by Liquidity Demander: This table looks at the market share at which a liquidity provider centrality group trades with a liquidity demander centrality group. The table reports the regression coefficients when the log market share of a liquidity provider to a liquidity demander is regressed on the liquidity demander's log number of liquidity providers and the centrality of the liquidity provider. The rows report the centrality of the liquidity demander, while the columns represent the liquidity provider. The coefficients show the regression coefficients of the liquidity demander and liquidity provider pair. I calculate the market shares as the share of trades a liquidity provider executes with a liquidity demander as a fraction of all the liquidity demander's trades. The sample of trades is constructed by looking at Streaming and RFQ trades with a trade size of at least 100,000 EUR. For Streaming trades, the trade size is the order size, not just the amount traded with the dealer in that trade. There are 10,438 observations, with each observation representing a liquidity provider and liquidity demander pair, and the unreported coefficient on log (*No. of Liquidity Provider*) is -1.04 with a standard error of 0.04

Log Market Share of Liquidity Provider by Liquidity Demander

	Top 10	11-40	41-115	116-250	251-450	451-1000	1000+
Top 10							
11-40	0.02	-0.9***	-0.76*	-0.9*		0.87	
	(0.35)	(0.28)	(0.42)	(0.47)		(0.67)	
41-115	-0.0	-0.81***	-1.12***	-1.39***	-1.6***	-0.73***	-2.24***
	(0.13)	(0.13)	(0.15)	(0.2)	(0.23)	(0.27)	(0.34)
116-250	-0.09	-0.59***	-0.83***	-1.15***	-1.32***	-1.09***	-1.25***
	(0.12)	(0.11)	(0.12)	(0.16)	(0.23)	(0.23)	(0.26)
251-450	-0.14	-0.53***	-0.73***	-0.78***	-1.07***	-0.47**	-0.95***
	(0.1)	(0.1)	(0.11)	(0.15)	(0.24)	(0.21)	(0.26)
451-1000	-0.18**	-0.35***	-0.53***	-0.57***	-0.79***	-0.4**	-0.72***
	(0.07)	(0.07)	(0.08)	(0.12)	(0.18)	(0.17)	(0.18)
1000+	-0.1	-0.15***	-0.1**	-0.15	-0.03	-0.05	-0.01
	(0.06)	(0.06)	(0.05)	(0.11)	(0.1)	(0.1)	(0.13)

Chapter 3

Robo-Advisers: Household Stock Market Participation and Investment Behavior

3.1 Introduction

In developed countries, stock market participation is surprisingly low. In Germany, for example, less than one-sixth of the population invests in stocks (Deutsches Aktien Insitut, 2019). In the US, the average "middle-aged" US family (a family led by an individual aged 41–60) holds more wealth in cars than in stocks (Ravikumar and Karson, 2018). Even high-income families invest far less in the stock market than might be expected. For example, only 40% of US households making an income between $100,000 and $200,000 even invest in stocks (Chien and Morris, 2017).

A lack of financial literacy is a key reason why households feel uncomfortable investing in the financial market (Deutsches Aktien Insitut, 2019) Until recently, reliance on the assistance of a professional financial adviser was the primary means of overcoming this obstacle. Nevertheless, this solution was out of reach for most households, as financial advisers commonly require high minimum investments (often 500,000 Euro) to compensate for the cost of human labor. However, this situation is beginning to change due to recent developments in FinTech, which have replaced human labor with machines in many financial services, thereby reducing costs[1]. However, such technologies have the potential to bring many personalized services, previously only available to high-net-worth households, to the average retail investor. One such service, at the focus of this study, is the Robo-Adviser—an automated system that invests clients' money according to academically vetted principles like Value-at-Risk or Strategic Asset Allocations while personalizing investment portfolios to each retail investor.

Hopes are that Robo-Advisers not only increase the diversification of retail investors, reduce the variance of their portfolio and attenuate biases that retail investors have when trading stocks, but also that more individuals start actively investing in the stock market. The Robo-Adviser's (frequent) trading is easily observable, as daily emails inform the retail investor of all the trades done by the Robo-Adviser. This straightforward observation of the Robo-Adviser's trading means that retail investors can easily get real-time feedback on what the Robo-Adviser is doing and how the investments are paying off, compared to traditional financial advice. This real-time feedback can influence the retail investors when investing money outside the Robo-Adviser, in the retail investor's active portfolio. For example, observing the lower volatility from diversification may have led the retail investor to diversify the portfolio she actively manages.

At the same time, the complete automation of Robo-Advisers can also lead to unwanted consequences. The retail investor's preferences are captured via a survey that might not adequately capture the retail investor's preferences and result in an undesirable portfolio allocation[2]. This problem is not unique to Robo-Advisers but instead arises whenever investing is

[1]There is a separate discussion of whether advisers deliver on the fees paid to them. Traditionally, these focused on human advisers with a strong focus on fees (Hackethal, Haliassos and Jappelli, 2012).

[2]The question of whether the Robo-Adviser chooses a suitable investment strategy for each retail investor is beyond the scope of this paper. While I do know the asset classes that a retail investor invests in, I do not know the weights of the individual securities in each asset class and, importantly, do not observe the survey responses to

delegated. Still, a multiple-choice survey may have more difficulty determining the preferences than a face-to-face conversation. Furthermore, the real-time information about the Robo-Adviser's trading may also have negative influences. For example, if retail investors replicate the Robo-Adviser's frequent trading, the retail investor's active portfolio's returns may reduce as fees rack up.

This paper aims to determine whether Robo-Advisers encourage new households to invest in financial markets and whether existing retail investors benefit from Robo-Advisers. To accomplish this aim, I examine how many investors in Robo-Advisers are new to the financial market and how many are not. For the latter, I determine whether the Robo-Adviser helps retail investors increase their returns across their whole portfolio, not just the portfolio invested with the Robo-Adviser.

While automation has made many services accessible to retail investors, the suitability of these services is yet unknown and remains to be studied. I acknowledge that the choice of a Robo-Adviser is not binary, with retail investors continuing to invest actively after investing with a Robo-Adviser. I use a proprietary data set from a bank cooperating with one of the largest Robo-Adviser in Germany. This data set's unique feature is that I observe retail investors' behavior using a Robo-Adviser and retail investors not using the Robo-Adviser and the Robo-Adviser's investment behavior. Furthermore, I follow the behavior of the users before and after using the Robo-Adviser.

Given the low participation of households in financial markets, especially in Germany, and the difficulty in increasing participation in financial markets, any increase in the participation in financial markets is beneficial. Examining a balanced panel of bank customers, including bank customers that invest with the Robo-Adviser (Robo-Adviser users) and bank customers that do not invest with the Robo-Adviser (control group), I find that a large portion of Robo-Adviser users had previously not participated in financial markets. In the year the Robo-Adviser was introduced, 2017, I observe that 35% of Robo-Adviser users have previously not participated in financial markets. For the control group, this fraction is only 5%. While I cannot rule out that these households own investment portfolios at other banks, I have reason to believe that these bank customers did not own a bank account at another bank.

Furthermore, my analysis shows that these new retail investors would not have participated in financial markets without introducing the Robo-Adviser. While participation in financial markets can lead to significant losses if invested carelessly, Robo-Advisers provide retail investors with high diversification, reducing the risks of investing. My central result is that the Robo-Adviser lives up to his promise and increases participation in financial markets by appealing to people that otherwise would not participate in financial markets.

While Robo-Advisers attract many households to become retail investors, investing is with a Robo-Adviser is not a binary choice. Instead, the majority of Robo-Adviser users invest actively on their own and with the Robo-Adviser. These Robo-Adviser users mainly use the Robo-

determine whether the allocation is suitable.

Adviser as a gimmick, investing close to the minimum required amount with the Robo-Adviser. Thus, the most pronounced effect the Robo-Adviser can have on these investors is changing how the Robo-Adviser users manage their active portfolio[3]. I compare Robo-Adviser users' active portfolio with control group retail investors' active portfolio by utilizing a difference-in-difference regression and controlling for observable characteristics before the Robo-Adviser introduction. In a robustness check, I find similar results using a propensity score matching based on characteristics before the introduction of the Robo-Adviser.

The Robo-Adviser investment differs from retail investor's investing by being more diversified and trading much more frequently. Looking at the margins of trading frequency and diversification of the actively managed portfolio, I find no evidence of any change in existing retail investors' investment behavior. I conclude that retail investors already participating in financial markets do not benefit from the Robo-Adviser.

I conclude that the Robo-Adviser increases participation in financial markets by attracting retail investors who would have otherwise not participated in financial markets, possibly due to their perceived lack of financial literacy. Furthermore, the initial investment amounts appear as a barrier to adoption. Many retail investors invest at or near the minimum investment amount, indicating they would ideally invest a lower amount. Furthermore, the Robo-Adviser does not seem beneficial for retail investors already participating in financial markets or ready to invest on their own in financial markets. However, given the limited usage of Robo-Adviser by existing retail investors, the gains of attracting new investors to financial markets outweigh the potential losses, in the form of higher fees, to the existing retail investors.

The remainder of the paper is structured as follows. The following section reviews current literature on Robo-Advisers and the investment behavior of households and retail investors. Section 3.3 describes the data. In Section 3.4, I describe the Robo-Adviser and how the Robo-Adviser's investing differs from retail investor's investing. Section 3.5 describes the structure of the analysis and creates a framework for how the Robo-Adviser is evaluated. Section 3.6 evaluates whether the Robo-Adviser benefits the retail investor. Section 3.7 performs robustness checks, before Section 3.8 concludes.

3.2 Literature Review

This paper studies the benefits of FinTech to consumers contributing to our understanding of the problems laid out in Goldstein, Jiang and Karolyi (2019). The closest paper to my paper is D'Acunto, Prabhala and Rossi (2019), which studies the effects of using software that provides additional advice when deciding how to trade. D'Acunto et al. (2019) study a very human-driven investment process, where the retail investor has to decide when to trade. For each trade,

[3]If the Robo-Adviser holds only a small part of the retail investor's wealth in the financial market, any benefit or cost on that part is small with respect to the whole portfolio. However, if the Robo-Adviser influences the behavior of the retail investor, the consequences affect the large actively managed portfolio.

she has to decide whether to use the software for suggestions and then decide to follow the software's suggestions. My study instead studies a fully automated investment process. The Robo-Adviser decides when and what to trade without any human intervention. Abraham, Schmukler and Tessada (2019) propose this definition of a Robo-Adviser and, to my knowledge, my study is the first to study Robo-Advisers as covered under their definition. D'Hondt, Winne, Ghysels and Raymond (2019) attempt to study a fully automated Robo-Adviser but rely on simulating a Robo-Adviser's behavior, rather than using a Robo-Adviser's actual trading, as this paper does. Furthermore, my study considers that using a Robo-Advisers is not binary, meaning that the Robo-Advisers manages either no or all investor investments. Instead, the investor can ask the Robo-Advisor to manage only a part of her portfolio. Thus, the investor continues to invest the other part of her portfolio actively. As I observe their management of the other part, I can examine how the retail investor's behavior changes after adopting the Robo-Adviser.

When analyzing the behavior of Robo-Advisers, I build on the literature on investment behavior to highlight pitfalls in how investors invest. This literature has shown retail investors suffer from underdiversification (Badarinza, Campbell and Ramadorai, 2016) and many behavioral biases (Odean (1998), Barber and Odean (2000) among others). While the existence of these biases is necessary to reduce these, there is much less work on reducing the biases in retail investor's portfolios, and much work seems to suggest that attempts to reduce these biases often fail (Bhattacharya, Hackethal, Kaesler, Loos and Meyer, 2012). While learning and experience often have limited effect in increasing retail investors performance (Koestner, Loos, Meyer and Hackethal, 2017), this paper explores whether delegating the investment decision to machines can increase the outcome for retail investors.

Previous work has shown that financial literacy is among the key determinants to investing in financial markets and has found that most people are financially illiterate. van Rooij, Lusardi and Alessie (2012) show that more financially literate households participate in financial markets and surveys confirm that the vast majority of people think that financial literacy is necessary to invest in financial markets (Deutsches Aktien Institut, 2019). For a complete overview on financial literacy, see for example Lusardi and Mitchell (2014). Furthermore, financial literacy does not only increase stock market participation but also once people are investing, higher financial literacy seems to correlate with reduced behavioral biases (Ateş, Coşkun, Şahin and Demircan, 2016; Grinblatt and Keloharju, 2000; Calvet, Campbell and Sodini, 2007) and overall seems to increase wealth creation (Lusardi and Mitchell, 2011). However, at the same time, Lusardi and Mitchell (2011) show that financial literacy is low in the United States and Jappelli (2010) finds Germany, the country studied in this study, has even lower financial literacy scores than the United States. Thus, increasing participation in financial markets through financial literacy seems at best a long-term effort.

Understanding the traditional importance of financial literacy for participation in financial markets (van Rooij et al., 2012) I ask whether there are alternative ways to increase participation in financial markets and whether FinTech can increase play a part in this, picking up suggestings by Goldstein et al. (2019). Especially since Gale, Harris and Levine (2012) argue that it is often difficult to determine whether financial literacy programs are helpful, so that alternative ways may be necessary to encourage participation in financial markets and ensure that retail investors achieve good returns.

3.3 Data Description

In this study, I worked with a large German retail bank. The bank specializes in low overhead costs, providing no bank branches and instead opting for a pure online banking platform. Thus, my sample consists of retail investors that are internet-affine. In October 2017, the bank started cooperating with a market leader in Robo-Advising services in Germany (based on amount managed), allowing the bank's customers to manage money with the Robo-Adviser. The bank's customers did not have to open another bank account, thus observing the Robo-Adviser's transactions within their bank account, although markedly separated. The service costs a yearly rate in the area of 0.5-0.8%, similar to other Robo-Advisers. These fees cover the service costs for both the bank and the Robo-Advising company. Only the expenses of the underlying ETFs that the Robo-Adviser buys. Signing up directly with the Robo-Adviser is arguably more difficult for the bank's customers since it requires opening up a new bank account at another bank, which is much smaller, on top of the investment mandate to the Robo-Adviser.

Furthermore, the fees paid when signing up directly with the Robo-Adviser are unchanged. Thus, bank customers who open an account with the Robo-Adviser are likely to do so via the bank. Furthermore, the bank takes pride in itself with a lean and straightforward product lineup that does not include an in-house active management service. Therefore, the bank does not suffer from an incentive problem to advise the product to its customers[4]. The minimum money to have managed by the Robo-Advising firm is 10,000 Euro.

For each year 2016, 2017, 2018, I observe all customers investing money with the Robo-Adviser in 2017. For these retail investors, I see the funds managed by the Robo-Adviser and the trades completed by the Robo-Adviser, and the money invested in financial markets under the retail investor's control. From now on, I call the funds not managed by the Robo-Adviser the retail investor's active portfolio, as the retail investor has active control about how to invest the money not under the Robo-Adviser's control. The portfolio allows me to see the asset classes invested in, the amount in each asset class, the number of securities held, the number of buys and sells each year, and the corresponding trading volumes.

I observe a random sample of around 10,000 retail investors that held at least 10,000 Euro in their account at the end of 2017 and traded at least once in 2017. These restrictions apply to have a comparable sample to the retail investors using the Robo-Adviser without removing the selection. 10,000 Euros invested was chosen as this is the minimum investment with

[4]As seen below, some banks closed their Robo-Advisers in 2018. However, they did have an in-house active management service that demanded larger fees than the Robo-Adviser.

the Robo-Adviser. I used one trade to ensure that the retail investors undertook at least one investment decision in 2017, as the retail investors using the Robo-Adviser did. I observe their active portfolio for these non-Robo-Adviser retail investors in the same detail as I do for the Robo-Adviser retail investors and observe each of these retail investors in 2016, 2017, 2018.

I assume that all bank customers observed in this sample have an investment portfolio with this bank. If they have an investment portfolio, i.e., if a retail investor is a bank customer and has an investment account, they have an investment account at the bank. This assumption allows me to conclude that retail investors with no portfolio at the bank in 2016 were not participating in financial markets.

I believe this assumption is reasonable as the bank provides comparatively cheaper brokerage services than other full-online banks with brokerage services. The bank charges both lower fixed commission and variable commission than comparable large banks in Germany, leading me to estimate that the trading cost is at least 10% lower at this bank. Therefore, opening a brokerage account with this bank is cheaper. Thus, even if the bank customers have additional bank accounts, they would have their portfolio with the given bank.

This assumption is especially important for the later Robo-Adviser users. At the time of the Robo-Adviser introduction, other major German banks introduced their own Robo-Adviser earlier in the same year. The other banks' Robo-Advisers invest similarly, and fee-wise are similar to the Robo-Adviser studied in this paper. Therefore, bank customers that were not trading with this bank but already had a portfolio at another bank would likely use the other banks Robo-Adviser. There they would have all their portfolios in one place and could easier monitor the effects of market movements on their portfolio. Thus, I believe that later Robo-Adviser users either had a portfolio at the bank or no portfolio at all.

3.4 Robo-Adviser Description

A Robo-Adviser is a product offered by a financial services company. The product is fully automated, requiring no human intervention to function (Abraham et al., 2019). Investing with a Robo-Adviser starts with the onboarding process. After filling out a survey, the retail investor's response gets translated into an investment strategy in an automated way. It continues with the investment, where the Robo-Adviser, an algorithm, automatically decides what and when to trade. This description of a Robo-Adviser, put forward by the World Bank in Abraham et al. (2019), differs markedly from the product studied in D'Acunto et al. (2019), where the retail investor has to choose to use the software for every trade, and then the retail investor decides to follow or deviate from the recommendation of the software.

The introduction of Robo-Advisers is very new to the German market. The oldest Robo-Adviser's exist since November 2013, with many being created only in 2016 or 2017. At the same time, more than 40 different firms are offering Robo-Advisers managing an estimated combined 2.8 bn Euro at the end of 2018 (Oliver Wyman, 2019). Despite some banks shutting their Robo-advising activity down in 2018[5], the assets under management have doubled in size each year between 2017 and 2020. Having such a young product makes the evaluation of Robo-Adviser difficult, therefore instead of looking at the return as the performance metric, I will study portfolio characteristics as discussed in more detail in the section 3.5.

3.4.1 Personalized Advice by a Robo-Adviser

Through the cooperation of the bank and firm offering the Robo-Adviser, the retail investor can open an account to manage his money by the Robo-Adviser through this bank directly[6]. The retail investor begins by filling out a survey that aims to understand the retail investor's investment preferences. Questions in the survey include the length of the investment horizon, literacy of financial products, income, wealth, liquid wealth, and investment aim.

After the retail investor has completed the survey, the retail investors' answers are translated into a Value-at-Risk which the Robo-Adviser targets. The Robo-Adviser then invests the money entrusted to it by the retail investor according to the Value-at-Risk corresponding to the survey and based on the Robo-Adviser's calculation of the different asset classes' moments. The Robo-Adviser invests in various asset classes exclusively through ETFs. ETFs cover the investment classes stocks, sovereign bonds, corporate bonds, covered bonds, real estate, and commodities and covering the regions, Europe, USA, Japan, Asia excluding Japan and Emerging Markets. All ETFs have a Total Expense Ratio between 0.07% and 0.40% per year. The Robo-Adviser uses the same ETF for an asset class across all Robo-Adviser users, so the same ETF always represents European stocks for all Robo-Adviser users.

After the Robo-Adviser allocates the money, the Robo-Adviser decides when to rebalance the portfolio and how to differ the asset allocation to maintain the Value-at-Risk target. The fees paid by the retail investor to the bank and the Robo-Adviser firm cover the costs of these trades.

Given that the portfolio is at the bank, the retail investor sees the Robo-Adviser portfolio separated from his portfolio but equally saliently. The retail investor receives emails for all trades done by the Robo-Adviser shortly after the trade execution[7].

---

[5]The banks shutting their Robo-Adviser down had an in-house active management service. Banks with an in-house active management might have a conflict of interest advising customers to open a Robo-adviser due to lower fees compared to active management. This conflict may have caused them to shut down their Robo-advising. (Finanz-Szene.com, Oct. 04, 2018, https://www.finanz-szene.de/digital-banking/immer-mehr-deutsche-banken-begraben-ihre-robo-projekte/)

[6]All Robo-Adviser's cooperate with a bank, but generally with a lesser-known smaller bank. The retail investor's money is kept at the bank, and the firm offering the Robo-Adviser has the authority to execute orders on the retail investor's account.

[7]Most often, the Robo-Adviser trades many securities when trading but trades relatively infrequently. So an investor is not swamped weekly by emails about the Robo-Advisor's trading.

3.4.2 Investment Behavior of a Robo-Adviser

To guide the analysis from now on, I compare the Robo-Adviser's behavior to the retail investor's behavior. As laid out below, I study portfolio characteristics, so only aspects where the Robo-Adviser differs from retail investors should lead to differently good portfolios.

The first difference comes directly from the investment style of the Robo-Adviser. By only investing in ETFs, the Robo-Adviser will manage a highly diversified portfolio and remove any idiosyncratic risk. Individual retail investors expose themselves much more to idiosyncratic risk. They often hold single stocks and no ETFs or Mutual Funds, with 20% of the Robo-Adviser Customers and almost 40% of the Control Group holding no ETFs or Mutual Funds[8]. Furthermore, holding only 10 securities on average, retail investors do not achieve diversification through holding individual stocks.

A second difference is the trading frequency by the Robo-Adviser. The Robo-Adviser trades much more frequently than the retail investors do, as Figure 1 shows, and comparing these 82.99 trades by the Robo-Adviser to the 21.85 trades by the Robo-Adviser Users in 2016 from Table 1. Implementing this large number of trades would be very costly for the retail investor, as they pay at least 4.95 Euro in fees to the bank per trade. Thus, trading this frequently would lead to at least an additional 302.64 Euro ($= 4.95 \times (82.99 - 21.85)$) in trading cost per year, three times the current (estimated) trading cost. The large number of Robo-Adviser trades come from rebalancing the portfolio, where every time the Robo-Adviser decides to sell a little from a few ETFs, while simultaneously a few other ETFs are bought a little, thereby slightly changing each constituent's weight in the portfolio. The Robo-Adviser achieves much lower costs through two steps. Firstly, the Robo-Adviser nets trading positions across Robo-Adviser users, leading to a small net trade compared to the many gross trades across users. Secondly, the Robo-Adviser has lower fees for trading by not paying a commission per trade. From the perspective of the Robo-Adviser user, the Robo-Adviser user only pays a fixed management fee, which includes the trading cost.

I will use these two differences in behavior between the Robo-Adviser and retail investors to study whether the Robo-Adviser benefits retail investors.

3.5 Structure of the Empirical Analysis

My data observes three portfolios, two for Robo-Adviser users and one for control group retail investors. For Robo-Adviser users, I observe the portfolio managed by the Robo-Adviser and the retail investor's active portfolio, which the Robo-Adviser user manages. While for control group retail investors, I observe their active portfolio. As I lay out below, my empirical results argue that the Robo-Adviser should mainly affect Robo-Adviser's portfolio through its effects on

[8]One might imagine Robo-Advisers being less prone to behavioral biases. However, while Robo-Advisers, through their non-human nature, cannot be prone to behavioral biases, data mining and other quantitative strategies may lead to similar adverse effects in trading.

the retail investor's active portfolio. Therefore, I compare the active portfolio of Robo-Adviser users with the active portfolio of the control group retail investors in a difference-in-difference regression.

Since both types of retail investors are bank customers, they have the choice to invest with the Robo-Adviser, but only Robo-Adviser users decide to do so. Not all retail investors are equally likely to invest with the Robo-Adviser. Therefore, Robo-Adviser users and control group retail investors are likely different. Therefore, I include retail investor's observable characteristics before the Robo-Adviser introduction in the regression. Thereby, I only compare similar Robo-Adviser users and the control group retail investors. I also performed propensity score matching on the pre-Robo-Adviser introduction retail investor's characteristics. Section 3.7 reports those results.

Next, I describe how I evaluate the three portfolios and describe under which conditions I compare which of the two portfolios with each other.

3.5.1 Evaluating Good and Bad Portfolios

Robo-Advisers are a very recent phenomenon. Thus, the time series of Robo-Adviser's behavior is very limited, spanning two years in my data. Therefore any observed returns will be driven by the time series specific return of an asset class, e.g., Bonds vs. Stocks or German stocks vs. US stocks. To remove this luck component, I will focus on the characteristics of a portfolio, such as diversification or biases. Retail investors display biases in their trading; for example, Barber and Odean (2000) show that trading very frequently does not increase returns but instead produces costs that can reduce the post-cost returns by 40%. In addition, many retail investors lack diversification in their portfolios. I will study whether the Robo-Adviser reduces the biases and increases diversification in the total portfolio, combining the Robo-Adviser's portfolio and the Robo-Adviser user's portfolio. If diversification increases or biases reduce, these changes allow the Robo-Adviser to achieve higher returns in the long run.

Focusing on the characteristics of a portfolio, like diversification or the number of trades in a period, replaces the realized return as the very noisy and business cycle biased estimate for the expected return[9], by characteristics of portfolios as the estimate for the expected [10]. This approach relies on determining the characteristics of a good portfolio. I lay out the characteristics describing a good portfolio shortly.

[9]Even for many retail investors, does the realized average return not represent the actual average return due to the effect of the market. The time series is too short to construct an unconditional average.

[10]I have the following framework in mind. An investor decides an investment strategy θ_i. Implementing the investment strategy θ_i produces a portfolio $P(\theta_i)$. I am interested in determining the return and volatility of the portfolio $P(\theta_i)$. Given the short time period, measuring the unconditional expected returns via the realized returns is impossible. However, θ_i gives me a portfolio $P(\theta_i)$. The properties of this portfolio can be much better mapped to expected returns or volatility. For example, diversification should lower volatility, or more frequent trading should reduce returns if a fixed cost is paid per trade. Thus analyzing the portfolio $P(\theta_i)$ instead of the return allows for a clearer inference on the expected return.

3.5.2 Determinants of a Good Portfolio

A good portfolio's characteristics might be debatable to some degree, but it clearly contains two following two characteristics. The first characteristic is diversification. In the absence of precise insider information, retail investors should not invest in only a few stocks but instead diversify by holding many stocks of a benchmark or an index product on the benchmark. The second characteristic is the cost of the portfolio, as good portfolios should have low fees. Fees include both fees to mutual fund managers but also trading fees. I focus on the latter to avoid discussing whether fees to mutual fund managers are justified or not (see, forBarras, Scaillet and Wermers (2010); Fama and French (2010)). Nevertheless, Barber and Odean (2000) support this view as they find that trading frequently does not increase the return but incurs substantial costs.

These two characteristics are arguably essential characteristics of a portfolio, as diversification allows investors to reduce their portfolio volatility by 50%. The average stock has the same return as the market but twice the volatility of the market. Transaction cost can quickly reduce the post-cost return of a portfolio by 40%, as Barber and Odean (2000) have shown. So, both of these characteristics have the potential to change the return of a portfolio in a meaningful way.

3.5.3 Robo-Advisor's (Potential) Influence on Portfolio Characteristics

When focusing on the characteristics of a good portfolio, it is essential to determine which characteristics to study and which characteristics can plausibly change. The first restriction will on be restricting what characteristics I will examine. I only focus on characteristics where the Robo-Adviser's portfolio differs from the retail investor's portfolio. Characteristics that the Robo-Adviser and the retail investor implement the same should lead to similar expected returns. Therefore, only among characteristics where the Robo-Adviser behaves differently from the retail investor should expected returns differ. Furthermore, when studying the actively managed part of the retail investor's portfolio, the characteristics that differ from the Robo-Adviser could change from observing the behavior of the Robo-Adviser. Thus, the Robo-Adviser may influence how the retail investor invests in their active portfolio.

As seen above, the most significant differences in the portfolio characteristics between the retail investor's active portfolio and the Robo-Adviser's portfolio are the trading frequency and the diversification of the Robo-Adviser. These two characteristics incidentally are also the characteristics that are unequivocally part of a good portfolio. In the analysis going forward, I, therefore, focus on these two characteristics.

3.5.4 Positive and Negative Changes Retail Investor's Portfolio Characteristics

The Robo-Adviser's portfolio is by construction diversified. The Robo-Adviser chooses only from a universe of 13 ETFs, which all have a sufficient number of holdings and are, therefore, by construction, highly diversified. On the other hand, retail investors choose not only from ETFs or Mutual Funds but also from individual stocks. Observing an increase in the ETF and Mutual Fund Share in the portfolio is likely associated with an increase in diversification, which I deem favorable. Furthermore, suppose the Robo-Adviser replaces individual stocks that the retail investor would have owned with ETFs, which the Robo-Adviser selected. In that case, I consider this replacement also as positive for the portfolio.

Robo-Adviser pays no commission[11] on an individual trade, I observe the Robo-Adviser trading much more frequently. On the other hand, the retail investor pays a commission for every trade she does, making trading much more costly. In Mifid II (introduced in January 2019), the EU requires banks and brokers to inform retail investors of trading costs pre-trade. However, this information was not readily available during my sample period. The fact that information on trading cost is later mandated suggests that retail investors did not understand or know the magnitude of trading cost. Observing the Robo-Adviser trading frequently and the academically vetted nature of the investment approach may lead the Robo-Adviser user to reasonably infer that frequent trading is beneficial. Instead, frequent trading creates higher trading costs for the retail investor. Therefore, observing an increasing trading frequency for the active portfolio of retail investors is considered negatively. In contrast, the Robo-Adviser's frequent trading is not viewed as unfavorable since there are no costs per trade to the investor[12]

3.5.5 Importance of Retail Investor's and Robo-Adviser's Portfolio to the Retail Investor

Finally, I explain how to compare the Robo-Adviser's portfolio with the retail investor's portfolio. Robo-Adviser users may have an active portfolio managed by themselves besides the portfolio with the Robo-Adviser, both of which constitute the retail investor's total portfolio. Changes in either portfolio's characteristics affect the characteristics of the combined portfolio.

Firstly, once investing with the Robo-Adviser, the fraction of money managed actively by the retail investor decreases. As the Robo-Adviser and the retail investor may differ in their investment behavior, i.e., their respective portfolios have different characteristics, the

[11] The retail investor faces no cost when the Robo-Advisor trades, except potentially the bid-ask spread, but the Robo-Adviser pays the bank more fees when completing more than 100 trades a year.

[12] One could argue that many trades by the Robo-Adviser create costs that the customers pay through the management fee of the Robo-Adviser. These costs could be passed on to the customer via a flat fee. However, the pass-through cannot be determined. In addition, the trading fees paid by the Robo-Adviser are much smaller than the retail investor pays, which means that the Robo-Adviser should be trading more frequently.

characteristic of the combined portfolio may change. I call changes due to the differences in characteristics of the Robo-Adviser's and the retail investor's active portfolio the composition channel. These changes arise since the difference in the composition of characteristics between the Robo-Adviser's and the retail investor's portfolio is the cause of these changes.

Secondly, investing with the Robo-Adviser may lead to changes in how the retail investor invests her money, changing the characteristics of her active portfolio. These changes arise from observing the investment behavior of the Robo-Adviser. These changes again affect the characteristics of the retail investor's combined portfolio. As these changes spillover from the Robo-Adviser to the retail investor, I call these changes the spillover channel.

Suppose the retail investor invests most of his wealth with the Robo-Adviser. In that case, the composition channel will determine the combined portfolio's characteristics. The spillover channel only affects the active portfolio, which only has a minor influence on the combined portfolio. However, suppose the retail investor invests only a little wealth with the Robo-Adviser. In that case, the spillover channel will likely dominate, as the composition channel from investing with the Robo-Adviser only exerts a minor influence on the combined portfolio.

For the composition channel, the comparison between the Robo-Adviser portfolio and the counterfactual active portfolio, i.e., the active portfolio of a comparable control group retail investor, is necessary to analyze. I compare the active retail investor's actual active portfolio to the counterfactual active portfolio to test the spillover channel, i.e., the control group retail investor's active portfolio. Thus, the analysis and comparison group rest upon how the retail investors invest with the Robo-Adviser, especially concerning the investment amount.

Therefore, I describe the users adopting the Robo-Adviser and what fraction of their wealth they invest with the Robo-Adviser.

3.6 Empirical Analysis

The sample consists of 9,551 Robo-Adviser users and 9,767 retail investors in the control group. A detailed description of the sample construction is in the appendix (Appendix C.1).

3.6.1 Robo-Adviser's Effect on the Participation of Retail Investors in Financial Markets

Households' participation in Financial Markets is low, and the hope is that Robo-Advisers will increase their participation in Financial Markets.

Understanding why a retail investor in 2017 might not have a portfolio in 2016 can be due to two reasons. First, the bank customer was a bank customer in 2016 but did not have a portfolio with the bank. Second, the bank customer was not yet a customer of the bank in 2016 and joined only in 2017. My analysis excludes the second group of new bank customers, as they might have held a portfolio at another bank. The bank provides trading services at lower costs than the closest competitors, and the closest competitors have also introduced Robo-Advisers. Therefore, I have no reason to believe that the first group was not active in financial markets in 2016 and are genuinely new traders (see Section 3.3).

Figure 2 shows that in 2016, roughly 5% of retail investors in both the control group and the group of Robo-Adviser users were new to financial markets. However, in 2017, Figure 3.4 shows that a substantial fraction of Robo-Adviser users (who are bank customers in 2016), namely 36.25%, are new participants in financial markets. Compared to the control group of retail investors, this group of retail investors newly participating in financial markets makes up only 7%. Thus, a much larger share of Robo-Adviser users is new to financial markets. They have previously not had an active portfolio with the bank than retail investors in the control group. Thus, showing that the Robo-Adviser attracts 3,462 new households to participate in financial markets newly.

While these investors are (likely) new investors to financial markets, the comparison does not reveal whether these investors would have invested in financial markets without introducing the Robo-Adviser. To address these concerns, I calculate whether the monthly number of new retail investors per month in 2017 changes after introducing the Robo-Adviser in October 2017[13]. Panel A of Table Table 3.1 compares the monthly number of new traders in the control group pre-and post-Robo-Adviser introduction. When looking at the accounts' opening throughout the year, I observe seasonal patterns with many retail investors opening their accounts at the beginning of the year. Thus, creating a bias predicting the number of new traders per month decreases at the end of the year. Therefore, finding no significant change in the number of new traders per month after introducing the Robo-Adviser underlines that these investors would have otherwise not participated in financial markets.

Panel B of Table 3.1 provides further support for these results. As there may be seasonal differences in the monthly number of new participants in financial markets across the year, Panel B includes the variable *Prob. to trade*$_i$ which is the probability that a new participant in financial markets starts trading in month i. *Prob. to trade*$_i$'s coefficient thus reports the number of new participants that year. Specification (1) in Table 2 shows the results from Panel A again. In Specification (2), I utilize only *Prob. to trade*$_i$ without a constant, which shows that the number of monthly new participants in financial markets dropped by an insignificant 0.467 traders, which represents a drop of less than $1\% (= \frac{0.467}{529/11})$ in the number of monthly new participants in financial markets. Specification (3) introduces a constant and continues to find very similar insignificant effects.

I, therefore, conclude that the Robo-Adviser does not motivate new active retail investors to invest with the Robo-Adviser instead of investing actively on their own, but rather motivates new households to participate in financial markets. So, the Robo-Adviser does not crowd out active investors. The Robo-Adviser, therefore, seems to attract many new retail investors to

[13]The post-Robo-Adviser introduction period are months October and November. The sample construction introduced no new investors in December 2017.

financial markets. It, therefore, seems to fulfill the hope of attracting especially new participants to financial markets, which is particularly crucial in a country like Germany, where only 16.2% of the population owns stocks through any participation. It, therefore, seems to contribute much to the overall welfare by encouraging new people to participate in financial markets.

3.6.2 Description of all Retail Investors Investing in 2016

To compare the groups before introducing the Robo-Adviser, Table 3.2 shows the summary statistics for the treatment group (Robo-Adviser users) and control group (retail investors in the control group) in 2016. It also outlines the differences between the groups. I restricted these groups to active retail investors in 2016. A retail investor is active if the retail investor has at least 2000 Euro in his portfolio at the end of the year, bought Securities valued at 2000 Euro, or sold Securities valued at 2000 Euro[14].

Table 3.2 depicts differences in the portfolios of the Robo-Adviser users and the retail investors in the control group. While the portfolio value is almost identical across the groups, the allocation is very different. Retail investors in the control group mainly invest in stocks, with 80.38% owning stocks and investing on average more than half their wealth in stocks, leading to a mean additional stock holding of 12,217.92 Euro Robo-Adviser users. Later Robo-Adviser users[15] instead are much more open to diversified products such as ETFs and Mutual Funds. They are 18.36 and 12.83 percentage points more likely to hold ETFs or Mutual Funds. The 12,217.92 Euro invested more into stocks by the retail investors in the control group is invested into ETFs and Mutual Funds by the later Robo-Adviser users. This difference in ETF and Mutual Fund ownership is the main difference. It shows that Robo-Adviser users are already more diversified[16] before the Robo-Adviser introduction. While academics have long taught the benefits of diversification, diversification only entered with the recent rise of passive investing through, for example, ETFs. Given only the recent rise of passive investing, many investors may not know or understand the benefits of diversification, thereby being put off by Robo-Advisers[17].

The Robo-Adviser users are also more likely to invest money into derivatives, which go by the name of "Zertifikate" in Germany. In contrast, the total value invested in them is generally tiny. Furthermore, Robo-Adviser users seem to less likely to be female, more often male, and attract more Joint Accounts[18]. Finally, the Robo-Adviser users trade more often by having three

[14]These restrictions removed mainly retail investor's with portfolio in 2016.

[15]I am looking at 2016, so these retail investors only later adopt the Robo-Adviser.

[16]The fact that the number of securities is the same across groups and only 10 highlights that the retail investors in the control group do not replicate the diversification by simply buying many stocks in the index.

[17]However, still at least 50% of control group retail investors own ETFs or Mutual Funds. Alternatively, ETFs and Mutual Funds are not just diversified but also delegate the investment decision. The higher likelihood of ETFs and Mutual Funds among Robo-Adviser users may therefore represent the higher propensity to delegate investing in general.

[18]Joint Accounts are the difference between the Percentage of Female and the Percentage of Males and makeup 22.77% of the control group accounts and 26.64% of the Robo-Adviser users.

more trades in 2016, which they realize by trading a smaller quantity per trade, as the total trade volume is the same.

3.6.3 Investment in the Robo-Adviser by Active Retail Investors in 2016

As outlined in Analysis Structure (Figure 3.2), to evaluate the Robo-Adviser's benefits to active retail investors, it is crucial to determine the amount invested with the Robo-Adviser. The spillover channel is important if and only if the Robo-Adviser manages a small part of the portfolio value. Likewise, the performance of the Robo-Adviser can only dominate potential behavioral changes (spillover channel) if the managed fraction by the Robo-Adviser is large.

As the Robo-Adviser requires a minimum investment of 10,000 Euro by the retail investor, the minimal mean fraction is already 28% (mean of $\frac{10{,}000}{10{,}000+\text{Portfolio Value 2017}}$).[19] Panel A of Figure 3.5 shows that the average fraction of money invested with the Robo-Adviser is 34.55%, indicating that retail investor invests only a small part of their portfolio with the Robo-Adviser. Panel B of Figure 3 further confirms these findings. It shows that in 2017 40% of investors invest the minimum amount with the Robo-Adviser, and 80% invest less than 20,000 Euro with the Robo-Adviser. In the analysis going forward, I focus on the actively managed portfolio and look for behavioral changes in the actively managed portfolio.

3.6.4 Effects from the Spillover Channel

3.6.4.1 Behavioral Changes in Active Portfolios of All Active Retail Investors

Since Robo-Adviser users invest close to the Robo-Adviser's minimal possible amount, the overall portfolio's characteristics mainly change through the behavior channel. Therefore, I compare the Robo-Adviser users' active portfolio with the control group retail investors' portfolio. I study the time series of the following characteristics: *Diversified*, whether the investor holds ETFs or Mutual Funds in their active portfolio; *Number of Asset Classes*, the number of assets classes an investor invests in, with the most common asset classes being Stocks, Bonds, ETFs and Mutual Funds; *ETF Share*, the percentage of the active portfolio invested in ETFs; *Mutual Fund Share*, the percentage of investments invested in Mutual Funds; and *Number of Trades*, the number of times a retail investor trades during a year. Figure 3.6 shows the time series of these characteristics for the Robo-Adviser Users and the control group.

[19] In unreported results, I find that the investment in the Robo-Adviser is additional money invested in financial markets. Most results support that all money invested with the Robo-Adviser is an additional investment. These results find that the active portfolio does not decrease in value, so there is no crowding out. However, even the result with the largest crowding out effect finds that the Net Investment (Buy Volume-Sell Volume) is only 10% of the total investment with the Robo-Adviser. Because even my largest estimates show that money invested with the Robo-Adviser is mainly additional investment into financial markets, I use $10,000$ + Portfolio Value 2017 as the total wealth in financial markets.

Figure 3.6 shows that there are likely no changes in the behavior across groups. Figure 3.6A shows that the gap in the holding of diversified assets is very much constant and is moving in lock-step with both rising and falling at the same time. While there are level differences, the difference in levels does not change. Similar behaviors can be seen in the number of asset classes (Figure 3.6B), ETF share (Figure 3.6C) and mutual fund share (Figure 3.6D). Where the last two show figures show either monotone increases in ownership or decreases in ownership. The only figure that shows a behavior that is not in lock-step is Figure 3.6E which shows the number of trades per retail investor. The figure shows that both Robo-Adviser users and control group retail investors increase their trading in 2017 and then decrease in 2018. However, the increase is greater for Robo-Adviser users, increasing their number of trades in 2017 by 4 trades. Control group retail investors increase their number of trades by only 2.5 over the same period. In 2018, both groups decreased their number of trades, but while Robo-Adviser users decreased their number of trades by 2.5 trades, the control group retail investors do so only by 1 trade. Thus, Figure 3.6 indicates that there are no behavioral changes upon adopting the Robo-Adviser, i.e., there is no spillover effect from the Robo-Adviser.

Figure 3.6 shows the group means of each group but does not control for the observable characteristics of each retail investor. I employ a difference in difference regression to control for the observable characteristics and compare a Robo-Adviser user to a comparable control group retail investor. Controlling for the observable characteristics of both the retail investor and the retail investor's portfolio in 2016, I treat the control group retail investor's active portfolio as the counterfactual active portfolio to the treated active portfolio of a Robo-Adviser user. I thus study the portfolio's difference after the adoption of the Robo-Adviser, as shown in Equation 3.1.

$$
\begin{aligned}
y_{i,t} = {} & \beta_1 \textit{Robo-Adviser User}_i + \beta_2 D(\textit{Robo-Adviser User} \times \geq 2017) \\
& + \beta_3 D(\textit{Robo-Adviser User} \times 2018) + \gamma_t + \delta X_{i,2016} + \epsilon_{i,t}
\end{aligned} \tag{3.1}
$$

In this setup, in the equation (Equation 3.1) above, I control for initial differences in behavior by adding the portfolio characteristics in 2016 ($X_{i,2016}$). I am interested in the coefficient β_3, which I define as the spillover impact of the Robo-Adviser. The coefficient β_3 is the incremental change of the 2018 difference between the Robo-Adviser Users and the Control Group. I look only at this incremental impact as the adoption of the Robo-Adviser happens only late in 2017. Thus, any "learning" from the Robo-Adviser likely occurs only in 2018. In addition, any unobserved factors that lead both to adoption in the Robo-Adviser and an increase in the dependent variable $y_{i,t}$ can contaminate the 2017 estimate. For example, a retail investor might want to diversify more and thus decide to buy both ETFs and adopt the Robo-Adviser, independent of his portfolio characteristics, in 2016. More generally, suppose a common unobserved factor leads to a change in the dependent variable and the Robo-Adviser's adoption. In that case, I should observe the changes simultaneously and thus see an impact on the 2017 coefficients but no further change on the 2018 coefficients.

Furthermore, learning from the Robo-Adviser takes time, and accordingly, changes in behavior induced by the Robo-Adviser should be observed only in 2018[20]. I cluster the standard errors at the individual level to account for time persistent individual effects.

Table 3.3 shows how retail investors change their active portfolio characteristics after investing with the Robo-Adviser. The studied characteristics are diversification and trading cost. Looking at diversification, Specification (1) and (2) in Table 3 precisely estimate that there is no change in the probability of owning diversified securities, like ETFs and mutual funds, or the number of asset classes a retail investor holds after investing money with the Robo-Adviser. Specifications (3) and (4) further confirm these results. Specification (3) shows that retail investors that invest money with the Robo-Adviser increase their investment share of ETFs by no more than retail investors increase their investment share of ETFs that do not invest in the Robo-Adviser. While Specification (4) finds a slight but significant decrease in the investment share of mutual funds in 2018, the coefficient is no longer significant when combined with the relative increase of the investment share in 2017 and beyond. Thus, this change is likely purely temporary, as many of these investment products have withdrawals only infrequently.

Looking at the trading frequency in Specifications (5) through (7), the results do not support the hypothesis that adopting the Robo-Adviser leads to increased trading by the retail investor, despite the Robo-Adviser trading on average 60 times more often than the average retail investor. Instead of increasing the trading frequency, the results indicate that adopting the Robo-Adviser leads to decreased trading frequency (-1.554 trades in 2018). As the money invested with the Robo-Adviser is almost exclusively additional money invested in financial markets, I would not expect a drop in the trading frequency. Closer inspection reveals that this drop in trading frequency results from an increase in the trading frequency in 2017 and beyond (1.580). Estimating an effect for only 2018 would result in a coefficient of only $0.026 (= 1.580 - 1.554)$. Thus, to the degree that Robo-Adviser users learn from the Robo-Adviser, these effects are purely temporary as they revert entirely in 2018 and only exist in 2017.

3.6.4.2 Behavioral Changes in Active Portfolios in Subgroups of Active Retail Investors

My results on the whole sample do not support the notion that retail investors change their investment behavior once investing with a Robo-Adviser. However, the whole group of Robo-Adviser users is on average well informed about diversification benefits, with more than 63% investing in Mutual Funds and 49% investing in ETFs. They already know that they should diversify, so the Robo-Adviser does not "teach" them that they should diversify. Likewise, most active retail investors have their active portfolio for a long time, with on average 8.37 years of

[20]Unreported results studied the effects on early and late adopters in 2017. Any positive comovement between the decision to adopt and the trading behavior should affect both groups similarly in 2017. Only aspects of learning should be stronger for the early adopters compared to the late adopters. If there is a positive effect from learning from the Robo-Adviser, the late adopters should have lower responses in 2017 but larger responses in 2018. I find that all incremental effects estimated on the late adopters are statistically indistinguishable from zero and tend to go in the opposite directions between 2017 and 2018.

experience, so they are less likely to learn given their large amount of experience. Therefore, changes in the trading frequency might also seem less likely, as they have experience. Therefore, in this section, I focus on subgroups that are more likely to learn from the Robo-Adviser. These are retail investors that are previously undiversified or inexperienced.

A retail investor is defined as an undiversified retail investor if they hold fewer than or equal to 5 securities and have no investment in either ETFs or Mutual Funds in 2016.[21] For this analysis, I continue to refer to undiversified retail investors (in 2016) as undiversified in 2017 and 2018, even if they own more than 5 securities, ETFs or Mutual Funds. Table 3.4 shows the summary statistic of the undiversified active retail investor in 2016. I observe a larger likelihood of control group retail investors being labeled undiversified with 16.80% of all control group retail investors compared to 12.45% for the Robo-Adviser users. This larger share in undiversified control group retail investors is likely higher if I removed the requirement to have invested 10,000 Euro in 2017. However, 10,000 Euro should be sufficient to diversify without incurring large trading costs due to the minimum commission per trade. This requirement of 10,000 Euro invested in 2017 likely adds to the increased average Portfolio Value of 22,229 Euros more. For the undiversified retail investors, the investment is almost exclusively in stocks, and Robo-Adviser users invest highly concentrated with almost 10,000 Euro invested in each security ($\frac{22{,}497}{2{,}42}$). I find that the later Robo-Adviser users are trading less frequently among this group than their control group retail investor counterparts. However, both groups trade frequently, turning over their whole portfolio once in 2016.

For the group of undiversified retail investors, I run the same regression as above in Equation 3.1, but add *Robo-Adviser User*$\times \geq t \times$ *Undiversified*-Dummies to study the additional effect on the undiversified.

$$\begin{aligned} y_{i,t} &= \beta_1 \textit{Robo-Adviser User}_i + \beta_2 D(\textit{Robo-Adviser User} \times \geq 2017) \\ &\quad + \beta_3 D(\textit{Robo-Adviser User} \times 2018) + \beta_4 D\left(\textit{Robo-Adviser User} \times \geq t \times \textit{Undiversified}\right) \\ &\quad + \beta_5 D\left(\geq t \times \textit{Undiversified}\right) + \gamma_t + \delta X_{i,2016} + \epsilon_{i,t} \end{aligned} \tag{3.2}$$

Table 3.5 show undiversified investors change their portfolio characteristics after adopting the Robo-Adviser and confirms the results of Table 3.3 for the broader population of Robo-Adviser Users. Concerning diversification, I observe a large and permanent increase in diversification in 2017. Undiversified Robo-Adviser users are $13 (= -0.0215 + 0.151)$ percentage points more likely to hold diversified securities compared to undiversified control group retail investors in 2017. The results indicate that an unobservable factor likely drives the decision

[21]I perform robustness among the number of securities and find that the results stay very similar. Allowing for investment into ETFs or Mutual Funds would require a specific threshold. However, I find that once a retail investor invests in either Mutual Funds or ETFs, the investment in these asset classes is often a substantial share of his wealth. Thus, I restrict myself from including them in this group, as any reasonable low threshold would only include relatively few additional retail investors.

to invest with the Robo-Adviser and the buying of diversified products, like ETFs and mutual funds, for example, learning about ETFs.

When looking at the marginal change in diversification in 2018 for the undiversified retail investors, I see that Robo-Adviser no longer significantly changes the diversification. The probability of holding diversified assets increases only an insignificant 2.04% (= 2.87% − 0.826%, t-stat 1.62) for undiversified Robo-Adviser users in 2018. Likewise, the marginal increase in mutual funds' share is insignificant compared to the undiversified control group retail investors in 2018 (0.456pp, t-stat 0.9).

At the same time, the coefficients in 2018 indicate more trading by the undiversified retail investors that invest with the Robo-Adviser in 2018, especially when compared with their pre-Robo-Adviser mean of 8.38 trades a year. However, these coefficients go opposite to the 2017 coefficients, and once I take the sum of these effects, I find that the total number of trades is unchanged, and the significance of the number of buys disappears. I thus find no evidence of any spillover effect for the trading frequency.[22]

Besides prior diversification, the Robo-Adviser likely influences less experienced retail investors more than experienced retail investors. With an average experience of 8.37 years, the Robo-Adviser likely has a much smaller effect than recent retail investors. Therefore, I also look at the subgroup of inexperienced retail investors. A retail investor is classified as inexperienced if the retail investor opened her portfolio only in 2016 and was thus only in his second year of trading in 2017.[23] As Table 3.6 shows the differences between the control group retail investors and the later Robo-Adviser users in 2016 are very similar to the differences between the groups for all active retail investors in 2016. The main differences being that inexperienced retail investors hold fewer securities than experienced retail investors and tend to invest in fewer asset classes. Not only the lower *# of Asset Classes* shows the investment in fewer asset classes, but also the consistently lower indicator variables for any asset class. I expected the portfolio value to be lower, as well as the shorter time being a customer. Finally, I observe that 30% of inexperienced control group retail investors be undiversified, while only 19% of inexperienced Robo-Adviser users are.

$$\begin{aligned} y_{i,t} = {} & \beta_1 \textit{Robo-Adviser User}_i + \beta_2 D(\textit{Robo-Adviser User} \times \geq 2017) \\ & + \beta_3 D(\textit{Robo-Adviser User} \times 2018) + \beta_4 D\left(\textit{Robo-Adviser User} \times \geq t \times \textit{Inexperienced}\right) \\ & + \beta_5 D\left(\geq t \times \textit{Inexperienced}\right) + \gamma_t + \delta X_{i,2016} + \epsilon_{i,t} \end{aligned} \tag{3.3}$$

In the analysis I again proceed by running the regression (Equation 3.3) and adding *Robo-Adviser User*$\times \geq t \times$ *Inexperienced*-Dummies. The results in Table 3.7 show that the

[22] At most, I would observe an increase in the number of buys with a corresponding decrease in the number of sells.

[23] I need to observe the 2016 portfolio characteristics in the Difference-in-Difference regression. Therefore, I restrict myself to retail investors with a portfolio at the end of 2016.

inexperienced do not change their investment into diversified assets. Neither can the hypothesis of no change in trading behavior be rejected, despite the number of asset classes invested in increasing marginally. Unreported results show more Robo-Adviser Users owning stocks (2.89pp larger Stock Ownership Ind. than inexperienced Control retail investors) leads to the increase in the number of asset classes held.

3.6.5 Conclusion on Existing Retail Investors' Combined Portfolio

The results show that investing with the Robo-Adviser does not lead to any changes in the retail investor's active portfolio. I find no changes to the investment characteristics even among the subsamples where the spillover channel is most likely due to poor previous investing or little prior knowledge. Table 3.2 shows that Robo-Adviser users invest significantly more in ETFs and Mutual Funds than control group retail investors, who instead invest that money into only a few individual stocks[24]. Given that the Robo-Adviser Users are already more likely to own ETFs and mutual funds, they probably understand that they should invest in diversified securities and are already doing so, so learning diversification from the Robo-Adviser is not necessary for them. They implemented investing in ETFs already beforehand. This result is in line with prior findings (Koestner et al., 2017) that retail investors have a hard time avoiding past mistakes.

The study of portfolio returns at the moment is very challenging, given the short time horizon. However, suppose the Robo-Adviser replaces an ETF or a mutual fund. In that case, the Robo-Adviser's performance should be compared to an index benchmark, where outperformance is generally difficult, as a large active management literature documents, for example, Fama and French (2010) or Barras et al. (2010).

3.7 Robustness Checks

In the previous section, I studied whether investing with a Robo-Adviser changes a retail investor's active portfolio characteristics. I used the characteristics of the retail investor and the characteristics of the retail investor's portfolio in 2016 as control variables to use all retail investors. In this section, I perform two robustness checks. First, I repeat the analysis, but instead of using the characteristics of the retail investors and their portfolios as control variables, I perform a propensity score matching to compare an investor to its closest match. Second, I repeat the analysis from Section 3.6 using postal code $\times$ year fixed effects.

[24]Table 3.2 also shows that Robo-Adviser users and control group retail investors hold the same number of securities, 10 on average, so that control group retail investors likely do not achieve diversification by buying many individual stocks.

3.7.1 Propensity Score Matching

In this section, I do a propensity score matching and find the nearest neighbor. Table 3.2 showed that retail investors that later invest with the Robo-Adviser differ from the retail investors that do not invest with the Robo-Adviser. I match each Robo-Adviser user to the most similar (nearest neighbor) control group retail investor to make these groups comparable. I calculate the propensity score using a logit regression. The variables used for control are the 2016 personal characteristics and the portfolio characteristics. The portfolio characteristics for investment decisions used for the logit regression are variations of the portfolio characteristics that are different from the Indicator used in the regression setup, such as Equation 3.1. Here the Indicator is replaced with the share invested in the asset class and the Euro-value invested in the asset class.

The results in Table 3.8 show that adopting the Robo-Adviser is particularly strongly predicted by not being female. Furthermore, the Robo-Adviser Users have higher trading volumes in 2016[25]. The results support that higher investments into ETFs and Mutual Funds increase the likelihood of adopting the Robo-Advisor.[26] Upon matching a control group retail investors with each Robo-Adviser user, Table 3.9 shows that the matching works well. Except for the probability of owning a stock or an ETF, all variables are matched very precisely and are indistinguishable across the groups, even if they are not part of the matching. The significant differences are, however, only statistically significant, but not economically significant, and the large differences observed in Table 3.2 no longer exist.

After I established that the propensity score matching via a logit regression is working well, I rerun the regression of Equation 3.1. However, the weighting difference arising from the oversampling of the treatment group (Robo-Adviser users) is no longer necessary. Instead, each observation is weighted equally. The results are displayed in Table 3.10 and they coincide with the results in Table 3.3. In Specification (1) through (4), I again find no changes in the diversification of a retail investor's portfolio once investing with the Robo-Adviser. Again, the mutual fund share decreases marginally but significantly in 2018, relative to 2017. However, when I study the cumulative change relative to 2016, the mutual fund share is again unchanged. Similarly, the coefficients on the trading frequency stay the same, and any drop in trading in 2018 is due to heightened trading in 2017, which subsequently reverses, as shown in Table 3.3. Thus, I confirm the previous results that the null of no change in the retail investor's active portfolio characteristics cannot be rejected.

[25] I report Sell Volume as a negative number thus the coefficient is positive in the absolute amount of selling.

[26] The inclusion of both investment shares, investment values, and the number of securities makes this challenging to be born out in the regression.

3.7.2 Postal Code and Time Fixed Effects

While the controls used in Equation 3.1 control for the observable differences in the retail investor and their portfolios, other important factors, such as the local economic environment, will differ across investors and affect their behavior. In order to control for those local effects that occur over time and affect only part of the population, I introduce *Postal Code* × *Year* fixed effects.[27]

The identification then comes from individuals within the same postal code and thus facing the same economic environment, one deciding to adopt the Robo-Adviser, while the other one decides not to adopt the Robo-Adviser. The results shown in Table 3.11 show that again the Robo-Adviser does not change the portfolio characteristics of the retail investor's active portfolio. I again estimate a precise 0 in the change probability of owning diversified securities or the number of asset classes. Furthermore, the investment into diversified securities is unchanged, with the share of money invested in ETFs or mutual funds again not changing. For the trading frequency, I find the same coefficients as in Table 3.3, with the cumulative effect after adopting the Robo-Investor being again a precise 0. Both robustness checks thus confirm that the previous finding that the Robo-Adviser does not change the portfolio characteristics of a retail investor's active portfolio.

3.8 Conclusion

Digitalization has expanded to Financial Services, making many products formerly exclusive to wealthy clients available to all retail investors. However, the suitability of these services is yet unknown. In this paper, I study one such service, namely financial advice, in the form of Robo-Advisers.

Using a novel data set that observes not only the trading by the Robo-Adviser but also the active trading of retail investors investing with the Robo-Adviser (Robo-Adviser users) and not investing with the Robo-Adviser (retail investors in the control group), I find that the Robo-Adviser serves a meaningful role by introducing 30% of its users to financial markets, who would have otherwise not participated in financial markets. These are households entirely new to financial markets, and without the Robo-Adviser, these households would not have invested actively on their own. Given the low stock market participation, especially in Germany, where only 16.2% of the population own stocks through any means of participation, increasing participation in the stock market is an essential service.

However, for retail investors already participating in financial markets, Robo-Advisers seem not to benefit them. Retail investors adopting the Robo-Adviser already own a significant amount of ETFs and Mutual Funds and, therefore, likely do not need the additional diversification

[27] I observe the retail investor's postal code in 2017, so if people move in 2016 or 2018, I do not observe the change in their postal code.

from Robo-Advisers. These retail investors currently invest only close to the minimum amount possible with the Robo-Adviser. The Robo-Adviser does not change their combined portfolio, i.e., the portfolio held by the Robo-Adviser and the active portfolio by the retail investor, in a meaningful way. These findings are supported by the Robo-Adviser not influencing the behavior of the retail investor. Given the much more frequent trading with the Robo-Adviser, one might expect the retail investors to trade more often after observing the Robo-Adviser. Still, I find no evidence of an increase in trading frequency. Given that these retail investors' diversification is already high, with more than 60% holding mutual funds and 50% holding ETFs, I find no increase in diversification among these groups. These findings hold even among subgroups most arguably most susceptible to these changes.

However, I find that the money invested with the Robo-Adviser constitutes at least 90% additional investment in financial markets, with at most 10% being money that would have been invested otherwise in the stock market.

Summing up, I conclude that the Robo-Adviser does not seem beneficial for retail investors who manage active portfolios independently. However, the Robo-Adviser increases participation in financial markets by attracting retail investors who would otherwise not have participated, possibly because of their perceived lack of financial literacy. These retail investors are not just investing in financial markets but immediately own very diversified portfolios. Attracting new investors to financial markets is particularly important given the low stock market participation in many countries, especially Germany, where only 16.2% of the population owns stocks. Nevertheless, the minimum investment amount required for such services may still be a barrier to adoption.

3.9 Chapter 3: Figures

Figure 3.1: Number of Trades by Robo-Adviser: This histogram shows the number of trades by the Robo-Adviser in 2018. The top graph shows the number of trades by the Robo-Adviser in 2018 across all Robo-Adviser users in 2017. The bottom graph restricts the sample to Robo-Adviser users who invest money with the Robo-Adviser in 2017 and at the end of 2018. I say a retail investor invests with the Robo-Adviser if her investment with the Robo-Adviser exceeds 1000 Euro.

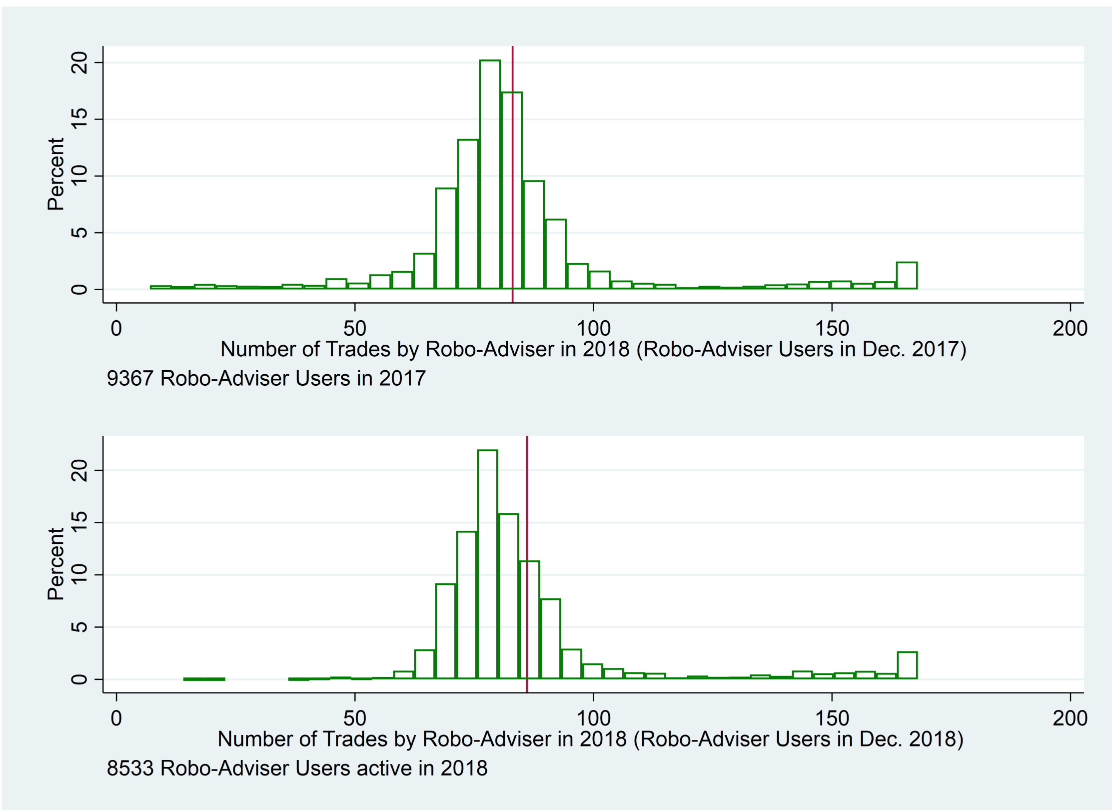


Figure 3.2: Example of Scenario when Spillover Channel Dominates: This figure depicts how a retail investor's portfolio using the Robo-Adviser and of a retail investor not using the Robo-Adviser change over time when the spillover channel dominates and thus the active portfolios are compared.

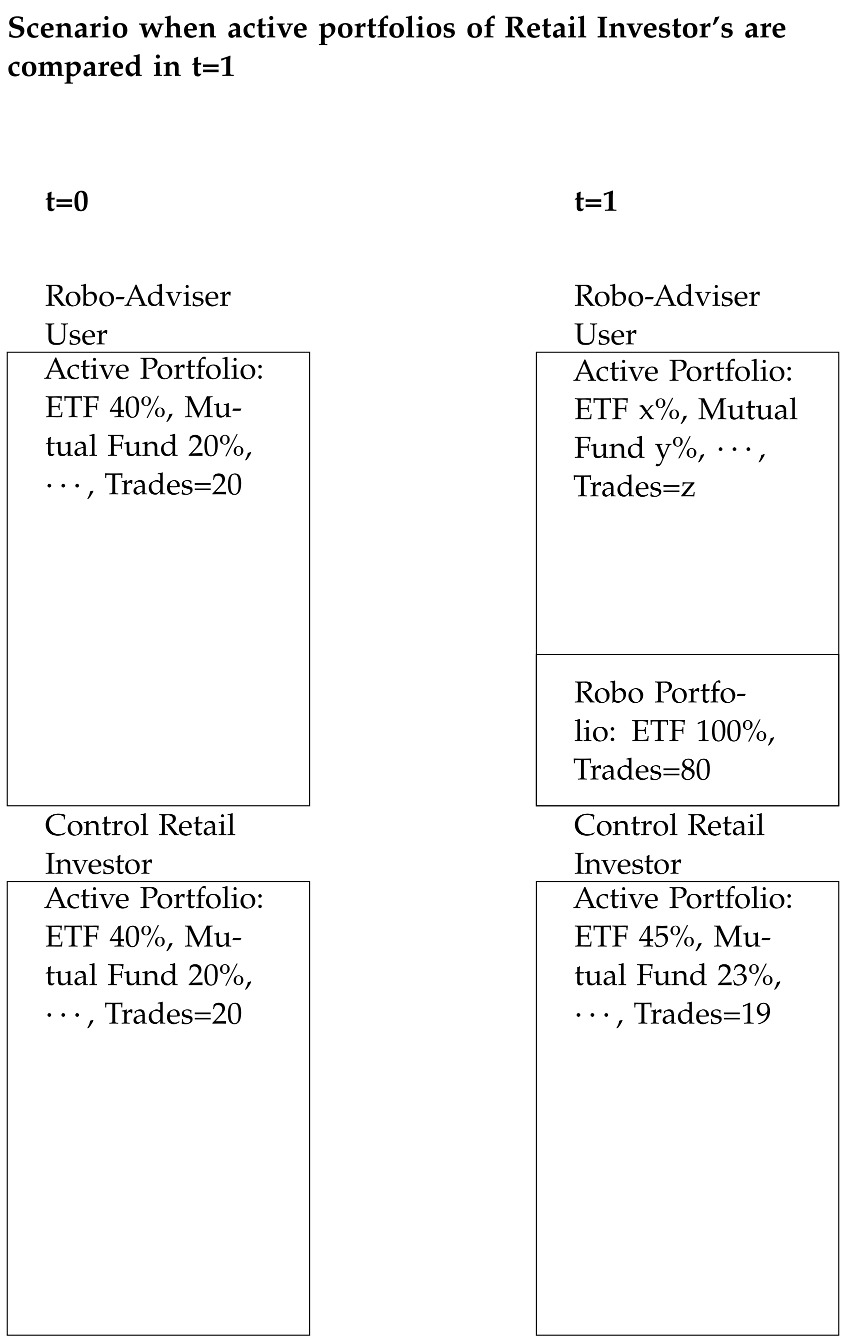


Figure 3.3: Example of Scenario when Composition Channel Dominates: This figure depicts how the portfolio of a retail investor using the Robo-Adviser and of a retail investor not using the Robo-Adviser change over time, when the composition channel dominates, and thus I compare the Robo-Adviser's portfolio to the portfolio of the control group retail investor.

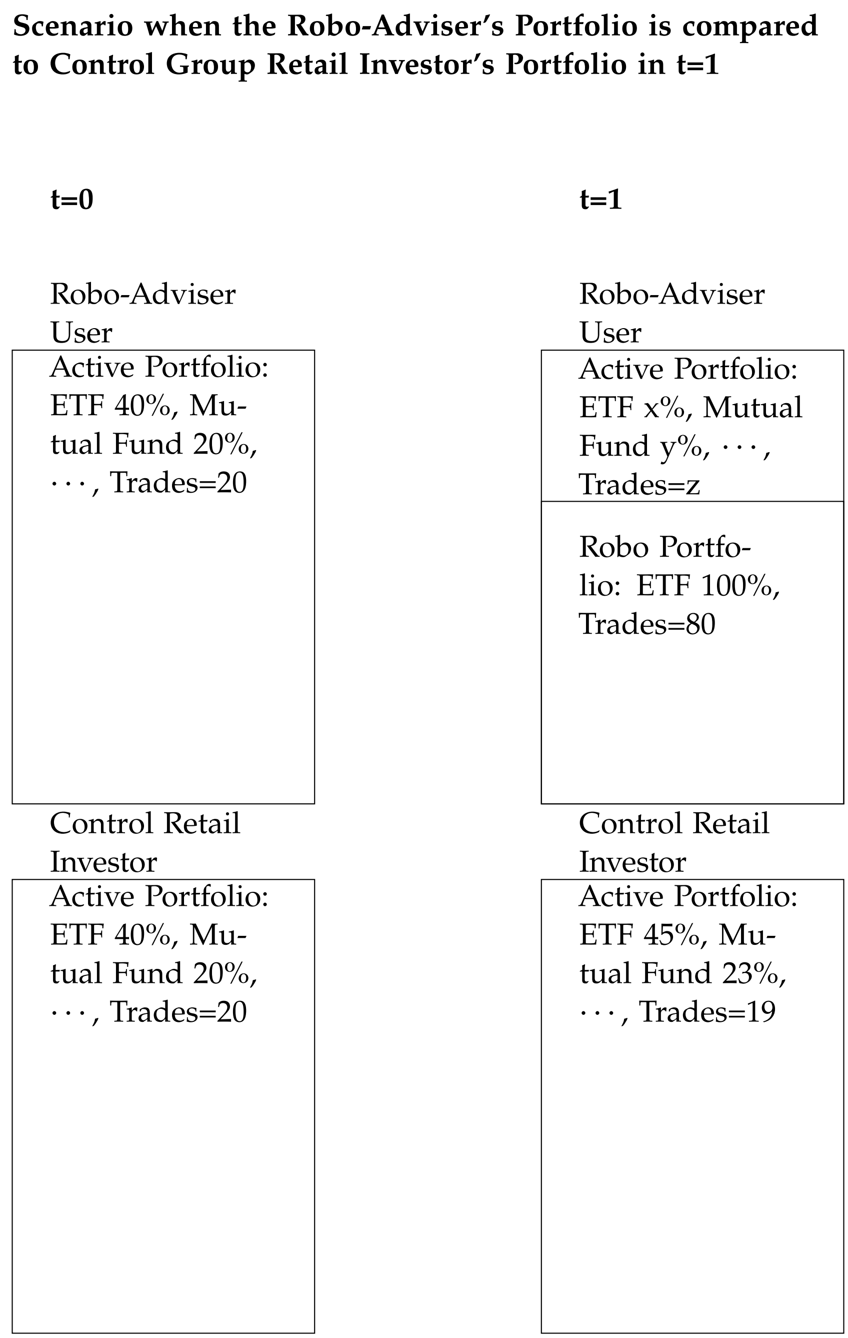


Figure 3.4: Share of Retail Investors Newly Participating in the Financial Market: This bar graph shows the share of new participants in financial markets for 2016 and 2017 in both the control group of retail investors and Robo-Adviser users. I classify a retail investor as a new participant in financial markets if the retail investor starts investing with either the Robo-Adviser or in an active portfolio for the first time that year.

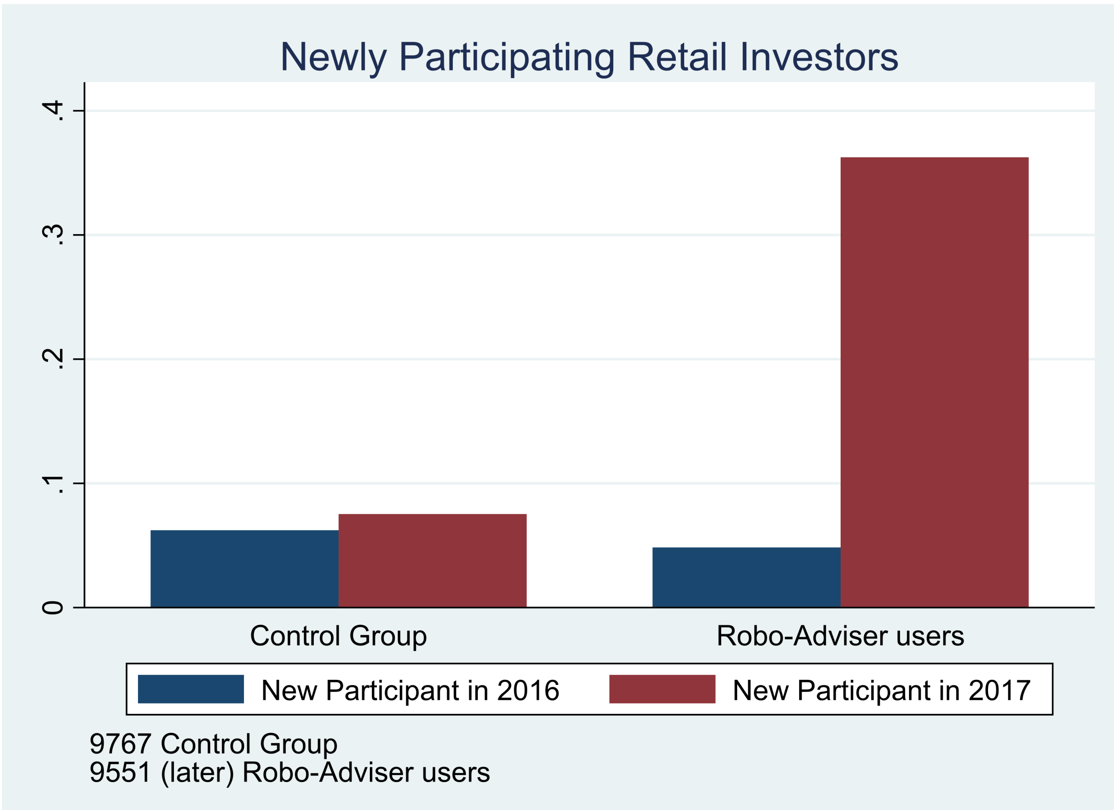
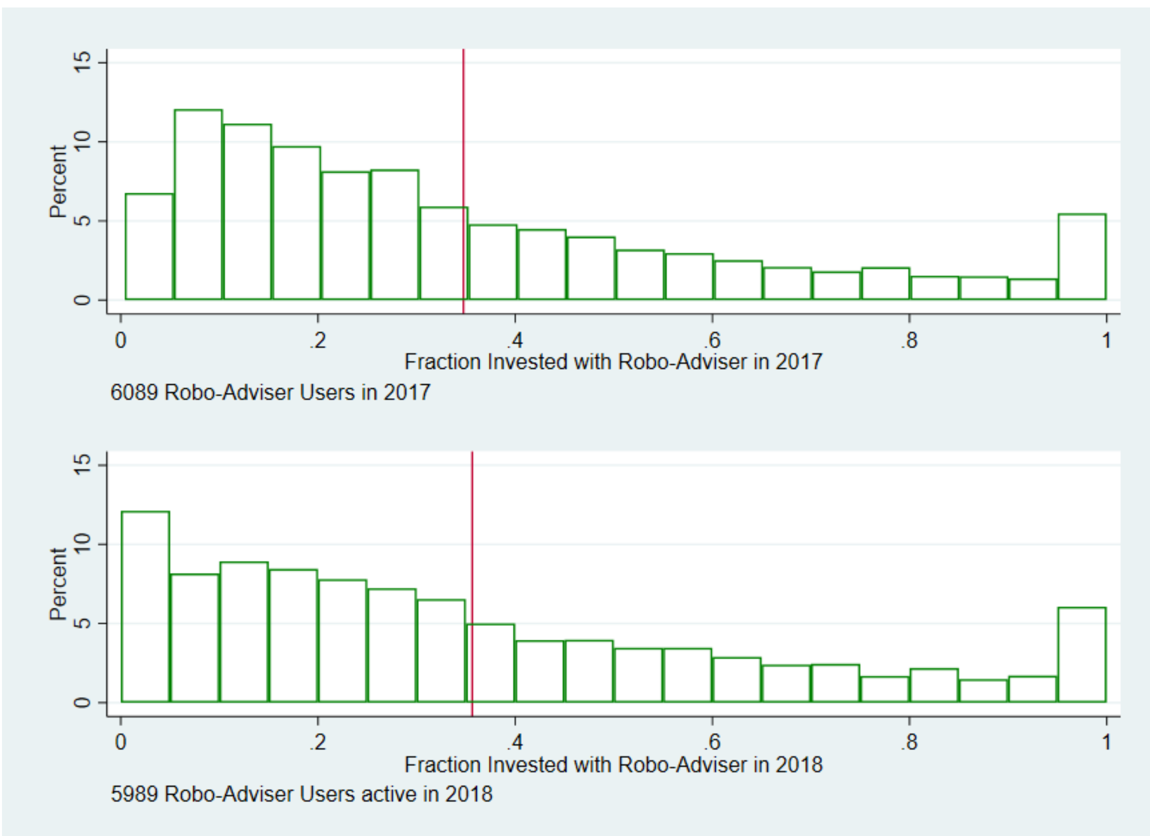
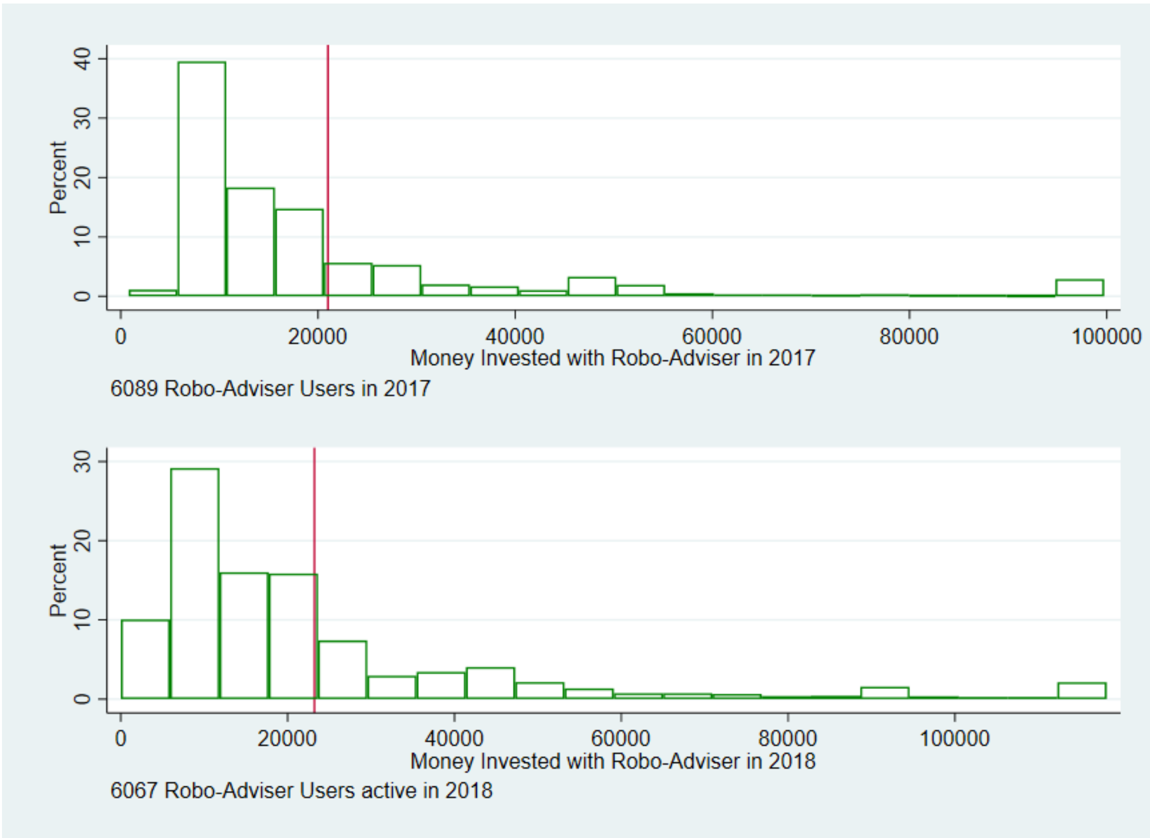


Figure 3.5: Distribution of Money Invested with the Robo-Adviser: The figure shows the distribution of the money invested with the Robo-Adviser in 2017 and 2018. **Panel A:** shows the fraction of the money managed by Robo-Adviser in comparison to the total investment in Financial Markets by the retail investor. In contrast, **Panel B:** shows the corresponding dollar amount distribution for investment amounts smaller than 100,000 Euro. Each Panel reports in the top half the fraction/the amount invested with the Robo-Adivser in 2017 and in the bottom half the fraction/the amount invested with the Robo-Adviser in 2018. The red line in each graph reports the mean (fraction of) money invested with the Robo-Adviser. The sample is constructed in 2017, by looking at active retail investors who are (later) Robo-Adviser users. In 2018, I further restrict the sample to only those users in the 2017 that still invest with the Robo-Adviser in 2018. I say a retail investor invests with the Robo-Adviser if her investment with the Robo-Adviser exceeds 1000 Euro.

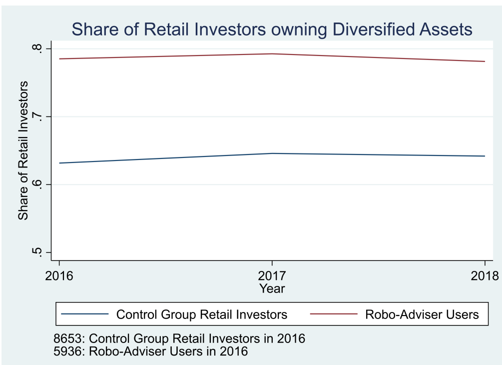


Panel A: Fraction of Money invested with the Robo-Adviser

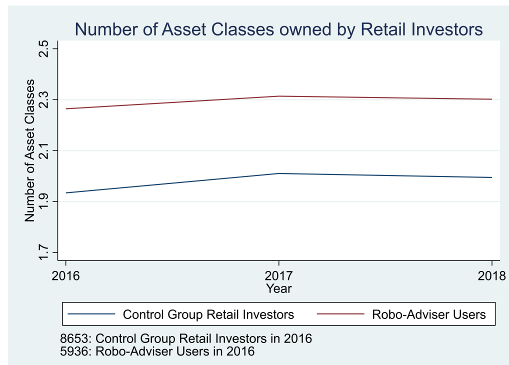


Panel B: Amount of Money invested with the Robo-Adviser

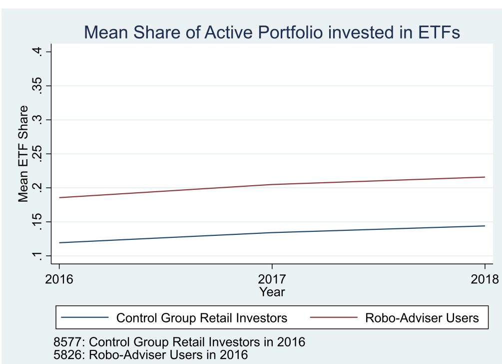
Figure 3.6: Time Series of Retail Investors' Active Portfolio Characteristics: The figure shows the time series of retail investors' active portfolio characteristics from 2016 to 2018. Retail Investors are Split into Robo-Adviser users and control group retail investors. **Panel** 3.6A**:** shows the share of investors that hold diversified assets, such as ETFs or Mutual Funds. **Panel** 3.6B**:** shows the number of asset classes the average retail investor in each group holds. **Panel** 3.6C and 3.6D**:** show the average share that an retail investors invests into either ETFs or Mutual Funds. And, **Panel** 3.6E**:** shows the number trades the average retail investor in each group executes each year.



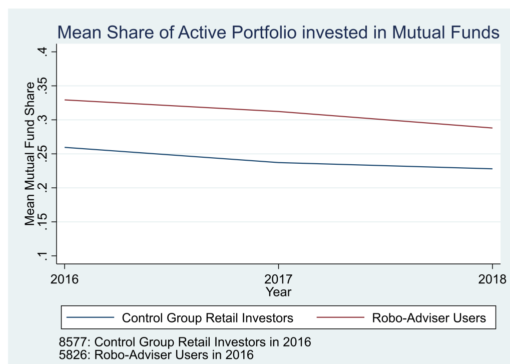
Panel A: Share of Diversified Retail Investors



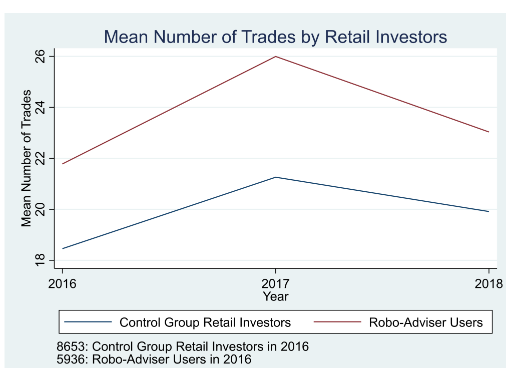
Panel B: Number of Asset Classes



Panel C: ETF Fund Share



Panel D: Mutual Fund Share



Panel E: Number of Trades

3.10 Chapter 3: Tables

Table 3.1: New Control Group Retail Investors each Month Pre- and Post-Robo Adviser Introduction: This table shows the difference in the number of traders pre- and post-Robo-Adviser introduction. **Panel A:** reports the difference in means pre- and post-Robo-Adviser introduction. In contrast, **Panel B:** controls once for the seasonal effects of starting to trade. Since new investors in December were excluded in the sample construction, I exclude the month of December is excluded in Panel A and Panel B. **Panel A:** Comparison in means. **Panel B:** This table is based on regressions to control for additional factors. Specification (1) tests whether the means differ pre- and post-Robo-Adviser introduction. Specification (2) and Specification (3) control for the difference in the probability of starting to trade each month (*Prob. to Trade*). *Prob. to Trade* as the average yearly distribution of new traders across a year. I calculate it as $\textit{Prob. to trade}_i = \frac{1}{T}\sum_{s=1}^{T} \frac{\textit{\#New Traders in month i in Year s}}{\textit{\#New Traders in Year s}}$.

Mean Number of New Traders each Month						
	Pre-Robo		Post-Robo		Difference	
	mean	sd	mean	sd	b	t
New Traders per Month	48.78	20.77	45.00	19.80	3.78	(0.23)
N	9		2		11	

Panel A: Comparison in Means

Regression of New Traders each Month			
	(1)	(2)	(3)
	New Traders per Month		
Post Robo-Introduction	-3.778	-0.467	-1.435
	(-0.29)	(-0.04)	(-0.12)
Prob. to trade		465.8***	408.9*
		(9.24)	(2.57)
Constant	48.78***		6.523
	(6.76)		(0.40)
N	11	11	11
R^2	0.00604	0.927	0.459

Panel B: This table tests whether the Robo-Adviser crowds out people from investing actively to instead adopting the Robo-Adviser, i.e., I test if the number of new traders starting trading drops after the introduction of the Robo-Adviser in year 2017.

Table 3.2: Summary Statistics of Active Retail Investors: This table shows the Summary Statistics for active retail investors in 2016. A retail investor is active if the retail investor had at least 2000 Euro in his portfolio at the end of 2016 or bought/sold 2000 Euro in stocks. It shows the summary statistics for Control Group retail investors (Control) and Robo-Adviser Users and the differences between the groups.

Summary Statistics of Active Retail Investors						
	Control		Robo-Adviser Users		Difference	
	mean	sd	mean	sd	b	t
Female (in %)	20.03	40.02	10.61	30.80	9.42***	(15.53)
Male (in %)	57.20	49.48	62.85	48.32	-5.65***	(-6.96)
Age	53.35	16.18	53.42	12.84	-0.07	(-0.28)
Years being Customer	12.23	4.91	12.89	4.84	-0.66***	(-8.18)
Portfolio Value (in 1000)	78.88	113.64	80.41	119.50	-1.53	(-0.80)
Portfolio Age	8.37	4.93	8.37	4.99	0.00	(0.01)
Stock Ind. (in %)	80.38	39.71	73.77	43.99	6.61***	(9.61)
Bond Ind. (in %)	13.23	33.88	16.57	37.18	-3.34***	(-5.71)
ETF Ind. (in %)	31.30	46.37	49.66	50.00	-18.36***	(-23.14)
Mutual Fund Ind. (in %)	50.81	50.00	63.64	48.11	-12.83***	(-15.71)
Levered Prod. Ind. (in %)	3.71	18.90	3.53	18.46	0.18	(0.57)
Derivative Ind. (in %)	10.09	30.12	15.91	36.58	-5.83***	(-10.69)
Other Security Ind. (in %)	3.11	17.36	2.96	16.94	00.16	(0.54)
Stock Value (in 1000)	48.75	93.64	36.53	78.85	12.22***	(8.37)
Bond Value (in 1000)	3.38	15.14	3.61	14.95	-233.61	(-0.94)
ETF Value (in 1000)	8.16	25.30	13.44	30.47	-5.28***	(-11.58)
Mutual Fund Value (in 1000)	17.38	38.54	24.80	46.65	-7.42***	(-10.66)
Levered Prod. Value	111.10	739.60	110.89	742.59	0.22	(0.02)
Derivative Value (in 1000)	1.03	4.81	1.84	6.44	-0.81***	(-8.83)
Other Security Value	70.79	562.24	76.92	586.29	-6.12	(-0.65)
# Buys	13.48	23.76	16.82	27.58	-3.34***	(-7.95)
# Trades	18.63	33.58	21.85	36.28	-3.23***	(-5.61)
# Sells	5.24	12.64	4.99	12.10	0.24	(1.17)
Buy Volume (in 1000)	39.75	101.22	39.63	93.60	0.12	(0.07)
Sell Volume (in 1000)	-30.80	92.20	-29.19	85.78	-1.60	(-1.08)
# Asset Classes	1.92	1.00	2.25	1.14	-0.33***	(-18.69)
# Securities	10.21	10.13	10.52	10.85	-0.31	(-1.77)
N	9,032		6,089		15,121	

Table 3.3: Retail Investors' Active Portfolio Characteristics' Change after Investing with a Robo-Adviser: This table shows how the portfolio characteristics of retail investors change after investing with a Robo-Adviser via a Difference-in-Difference regression. It compares the actively managed portfolio of the Robo-Adviser Users with the (actively managed) portfolio of Non-Robo-Adviser Users from 2016 (before the Robo-Advisor introduction) to 2018. It includes only retail investors with an active portfolio in 2016. The dependent variables are: *Diversified Ind.* is an indicator equal to 1 if the retail investor holds ETFs or Mutual Funds in his (active) portfolio. Asset Classes count the number of asset classes in the retail investor's portfolio. *ETF Share*, *MF Share* report the fraction of the active portfolio held in Mutual Funds or ETFs. *Trades*, *Buys*, *Sells* reports the number of trades/buys/sells during that year by the retail investor. The Weighting Correction refers to the Correction used to adjust for the Robo-Advisor Users' Oversampling relative to the Non-Robo-Advisor Users. The panel is unbalanced with investors leaving in 2018. The control variables include personal characteristics: Sex, Joint Account, Age and portfolio characteristics: Years as Customer, Years since opening the portfolio and the 2016 investment characteristics portfolio value, stock dummy, bond dummy, ETF dummy, Mutual Funds Dummy, Levered Product Dummy, Derivative Dummy, Buy Volume, Sell Volume, Number of Asset Classes, Number of Securities and Turnover. The standard errors cluster at the individual level.

Change in Retail Investors Active Portfolio Characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Diversified Ind.	Asset Classes	ETF Share	MF Share	Trades	Buys	Sells
$D(Robo)$	0.000980	0.0315***	0.00350	0.0199***	0.138	0.129	-0.0331
	(0.22)	(9.63)	(0.98)	(4.89)	(1.05)	(1.49)	(-0.64)
$D(Robo \times \geq 2017)$	-0.00756	-0.0229*	0.00446	0.00675*	1.580***	0.907***	0.661***
	(-1.57)	(-2.00)	(1.60)	(2.46)	(4.15)	(3.34)	(4.23)
$D(Robo \times 2018)$	-0.00601	0.00643	0.00210	-0.0129***	-1.554***	-1.204***	-0.306
	(-1.52)	(0.69)	(0.91)	(-5.54)	(-4.06)	(-4.44)	(-1.95)
N	43,162	43,162	42,810	42,810	43,162	43,162	43,162
R^2	0.646	0.774	0.521	0.591	0.706	0.716	0.662
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighting Correction	Yes	Yes	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

Table 3.4: Summary Statistics of Undiversified Retail Investors: This table reports the summary statistics for undiversified retail investors. An retail investor is called undiversified if the retail investor owned less than 5 assets in 2016 and no Mutual Funds or ETFs.

Summary Statistics of Undiversified Retail Investors						
	Control		Robo-Adviser Users		Difference	
	mean	sd	mean	sd	b	t
Female (in %)	19.60	39.71	12.04	32.57	7.56***	(4.47)
Male (in %)	59.90	49.03	63.06	48.30	-3.15	(-1.43)
Age	51.89	16.40	51.85	13.65	0.03	(0.04)
Years being Customer	10.98	5.07	12.42	4.74	-1.44***	(-6.41)
Portfolio Value (in 1000)	42.67	74.22	22.50	45.963	20.17***	(6.76)
Portfolio Age	6.86	4.80	7.28	4.87	-0.42	(-1.92)
Stock Ind. (in %)	93.95	23.85	89.85	30.22	4.10***	(3.46)
Bond Ind. (in %)	8.87	28.44	11.23	31.60	-2.36	(-1.77)
Levered Prod. Ind. (in %)	1.93	13.75	2.57	15.84	-0.65	(-0.99)
Derivative Ind. (in %)	4.40	20.52	7.98	27.12	-3.58***	(-3.45)
Other Security Ind. (in %)	1.38	11.65	1.62	12.65	-0.25	(-0.46)
Stock Value (in 1000)	39.49	74.44	19.31	43.94	20.17***	(6.79)
Bond Value (in 1000)	2.48	12.07	2.12	10.17	0.37	(0.71)
Levered Prod. Value	52.53	478.91	88.66	673.84	-36.12	(-1.45)
Derivative Value (in 1000)	0.61	3.89	0.91	4.66	-0.30	(-1.58)
Other Security Value	36.92	419.87	66.84	550.45	-29.92	(-1.42)
# Buys	7.54	17.26	4.84	13.69	2.71***	(3.71)
# Trades	12.58	28.19	8.38	22.32	4.20***	(3.53)
# Sells	5.12	11.93	3.59	9.89	1.53**	(3.00)
Buy Volume (in 1000)	49.31	120.96	28.14	93.20	21.17***	(4.17)
Sell Volume (in 1000)	44.27	117.11	27.93	90.93	16.34***	(3.32)
# Asset Classes	1.11	0.33	1.13	0.38	-0.03	(-1.75)
# Securities	2.89	1.39	2.42	1.37	0.47***	(7.55)
N	1,454		739		2,193	

Table 3.5: Changes to Undiversified Retail Investor's Portfolio Characteristics after Investing with the Robo-Adviser: This table shows the regression results of a Difference-in-Difference regression. It compares the actively managed portfolio of the Robo-Adviser Users with the (actively managed) portfolio of Non-Robo-Adviser Users from 2016 (before the Robo-Advisor introduction) to 2018. It included only retail investors with an active portfolio in 2016. A retail investor is classified as undiversified (Undiv.), if in 2016 the investor held fewer or equal to 5 assets, with none of the assets being an ETF or a Mutual Fund. The dependent variables are: *Diversified Ind.* is an indicator being 1 if the retail investor holds ETFs or Mutual Funds in his (active) portfolio. *Asset Classes* counts the number of asset classes in the retail investor's portfolio. *ETF Share*, *MF Share* report the fraction of the active portfolio held in Mutual Funds or ETFs. *Trades*, *Buys*, *Sells* reports the number of trades/buys/sells during that year by the retail investor. The Weighting Correction refers to the Correction used to adjust for the Oversampling of the Robo-Advisor Users relative to the Non-Robo-Advisor Users. The panel is unbalanced with investors leaving in 2018. The control variables include personal characteristics: Sex, Joint Account, Age and portfolio characteristics: Years as Customer, Years since opening the portfolio and the 2016 investment characteristics portfolio value, stock dummy, bond dummy, ETF dummy, Mutual Funds Dummy, Levered Product Dummy, Derivative Dummy, Buy Volume, Sell Volume, Number of Asset Classes, Number of Securities and Turnover. The standard errors are clustered at the individual level.

	Change in Retail Investors Active Portfolio Characteristics						
	(1) Diversified Ind.	(2) Asset Classes	(3) ETF Share	(4) MF Share	(5) Trades	(6) Buys	(7) Sells
$D(Robo \times \geq 2017)$	-0.0215***	-0.0293*	-0.00128	0.00436	2.019***	1.118***	0.893***
	(-4.67)	(-2.38)	(-0.43)	(1.45)	(4.95)	(3.80)	(5.33)
$D(Robo \times 2018)$	-0.00826*	0.00215	0.000958	-0.0146***	-2.007***	-1.519***	-0.431*
	(-2.00)	(0.21)	(0.39)	(-5.69)	(-4.93)	(-5.18)	(-2.57)
$D(Robo \times \text{Undiv.} \times \geq 2017)$	0.151***	0.0895**	0.0503***	0.0311***	-2.856**	-1.136	-1.724***
	(8.18)	(2.69)	(5.67)	(4.43)	(-2.58)	(-1.55)	(-3.84)
$D(Robo \times \text{Undiv.} \times 2018)$	0.0287*	0.0441	0.00794	0.0192***	3.419**	2.408**	0.926*
	(2.17)	(1.72)	(1.14)	(3.38)	(3.01)	(3.22)	(1.96)
N	43,162	43,162	42,810	42,810	43,162	43,162	43,162
R^2	0.668	0.775	0.528	0.594	0.706	0.716	0.663
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighting Correction	Yes	Yes	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

Table 3.6: Summary Statistics of Inexperienced Retail Investors: This table reports the Summary Statistics for inexperienced retail investors. An retail investor is classified as inexperienced if the retail investor began trading in 2016.

Summary Statistics of Inexperienced Retail Investors						
	Control		Robo-Adviser Users		Difference	
	mean	sd	mean	sd	b	t
Female (in %)	25.39	43.56	13.04	33.71	12.35***	(5.17)
Male (in %)	51.25	50.02	.61.49	48.71	-10.24***	(-3.44)
Age	49.81	16.91	51.13	13.58	-1.32	(-1.41)
Years being Customer	7.42	5.99	8.65	6.45	-1.23**	(-3.30)
Portfolio Value (in 1000)	39.71	66.85	35923.20	65.47	3.79	(0.95)
Portfolio Age	0.59	0.30	0.51	0.32	0.09***	(4.55)
Stock Ind. (in %)	66.82	47.12	54.04	49.89	12.79***	(4.39)
Bond Ind. (in %)	9.03	28.69	11.59	32.05	-2.56	(-1.41)
ETF Ind. (in %)	33.64	0.4729	49.69	50.05	-16.04***	(-5.49)
Mutual Fund Ind. (in %)	31.93	46.66	43.69	49.65	-11.75***	(-4.07)
Levered Prod. Ind. (in %)	1.56	12.39	3.11	17.36	-1.55	(-1.74)
Derivative Ind. (in %)	5.30	22.41	5.80	23.39	-0.50	(-0.36)
Other Security Ind. (in %)	1.25	11.10	1.45	11.96	-0.20	(-0.29)
Stock Value (in 1000)	22.05	49.37	15.07	46.88	6.98*	(2.40)
Bond Value (in 1000)	2.13	11.12	1.54	8.02	0.59	(0.99)
ETF Value (in 1000)	6.86	21.81	9.13	22.55	-2.27	(-1.71)
Mutual Fund Value (in 1000)	8.09	27.16	9.53	26.03	-1.43	(-0.89)
Levered Prod. Value	37.73	405.50	102.60	730.75	-64.87	(-1.89)
Derivative Value (in 1000)	0.51	3.42	0.52	3.17	-0.01	(-0.05)
Other Security Value	34.57	374.95	32.69	385.24	1.88	(0.08)
# Buys	10.53	16.56	8.86	16.24	1.67	(1.69)
# Trades	13.64	23.68	11.12	22.42	2.52	(1.80)
# Sells	3.14	8.74	2.22	7.59	0.91	(1.83)
Buy Volume (in 1000)	36.00	80.19	27.43	65.61	8.57	(1.91)
Sell Volume (in 1000)	17.62	63.15	10.59	49.88	7.04*	(2.02)
# Asset Classes	1.49	0.84	1.69	0.92	-0.20***	(-3.76)
# Securities	5.99	6.39	5.44	6.15	0.55	(1.45)
Undiversified (in 2016)	0.30	0.46	0.19	0.39	0.12***	(4.43)
N	642		483		1,125	

Table 3.7: Changes to Inexperienced Retail Investor's Portfolio Characteristics after Investing with the Robo-Adviser: This table shows the regression results of a Difference-in-Difference regression. It compares the actively managed portfolio of the Robo-Adviser Users with the (actively managed) portfolio of Non-Robo-Adviser Users from 2016 (before the Robo-Advisor introduction) to 2018. It included only retail investors with an active portfolio in 2016. A retail investor is classified as inexperienced (Inexp.) if the investor opened an account in 2016. The dependent variables are: *Diversified Ind.* is an indicator being 1 if the retail investor holds ETFs or Mutual Funds in his (active) portfolio. *Asset Classes* counts the number of asset classes in the retail investor's portfolio. *ETF Share*, *MF Share* report the fraction of the active portfolio held in Mutual Funds or ETFs. *Trades*, *Buys*, *Sells* reports the number of trades/buys/sells during that year by the retail investor. The Weighting Correction refers to the Correction used to adjust for the Oversampling of the Robo-Advisor Users relative to the Non-Robo-Advisor Users. The panel is unbalanced with investors leaving in 2018. The control variables include personal characteristics: Sex, Joint Account, Age and portfolio characteristics: Years as Customer, Years since opening the portfolio and the 2016 investment characteristics portfolio value, stock dummy, bond dummy, ETF dummy, Mutual Funds Dummy, Levered Product Dummy, Derivative Dummy, Buy Volume, Sell Volume, Number of Asset Classes, Number of Securities and Turnover. The standard errors are clustered at the individual level.

Change in Retail Investors Active Portfolio Characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Diversified Ind.	Asset Classes	ETF Share	MF Share	Trades	Buys	Sells
$D(Robo \times \geq 2017)$	-0.00567	-0.0266*	0.00657*	0.00650*	1.427***	0.803**	0.605***
	(-1.16)	(-2.27)	(2.36)	(2.32)	(3.70)	(2.91)	(3.80)
$D(Robo \times 2018)$	-0.00677	0.000141	0.00313	-0.0128***	-1.684***	-1.331***	-0.313
	(-1.68)	(0.01)	(1.33)	(-5.35)	(-4.23)	(-4.72)	(-1.90)
$D(Robo \times \text{Inexp.} \times \geq 2017)$	-0.0273	0.0385	-0.0264	0.00275	1.622	1.114	0.589
	(-1.21)	(0.77)	(-1.81)	(0.22)	(0.91)	(0.90)	(0.84)
$D(Robo \times \text{Inexp.} \times 2018)$	0.00911	0.0761*	-0.0133	-0.00132	1.823	1.766	0.0872
	(0.53)	(2.08)	(-1.29)	(-0.13)	(1.28)	(1.72)	(0.15)
N	43,162	43,162	42,810	42,810	43,162	43,162	43,162
R^2	0.646	0.774	0.521	0.591	0.706	0.716	0.663
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighting Correction	Yes	Yes	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

Table 3.8: Logit Regression Predicting the Decision to Invest with a Robo-Adviser: This table reports the result of the logit regression used for the propensity score matching. The Weighting Correction refers to the Correction used to adjust for the Oversampling of the Robo-Advisor Users relative to the Non-Robo-Advisor Users. The control variables include personal characteristics: Sex, Joint Account, Age and portfolio characteristics: Years as Customer, Years since opening the portfolio and the 2016 investment characteristics portfolio value, stock dummy, bond dummy, ETF dummy, Mutual Funds Dummy, Levered Product Dummy, Derivative Dummy, Buy Volume, Sell Volume, Number of Asset Classes, Number of Securities and Turnover. The control variables are the variables used for matching.

Predicting Investment with a Robo-Adviser	
	(1)
Male	0.0208
	(0.46)
Female	-0.732***
	(-11.84)
Years being Customer	0.0407***
	(8.02)
Age	0.00150
	(1.21)
Portfolio Age	-0.0338***
	(-6.74)
Portfolio Value	-0.0000234
	(-0.59)
Stock Share	-1.776
	(-1.59)
Bond Share	-1.412
	(-1.25)
ETF Share	-0.897
	(-0.80)
Mutual Fund Share	-1.191
	(-1.07)
Levered Prod. Share	-1.968
	(-1.53)
Derivative Share	-0.880
	(-0.77)

Continued on next page

Predicting Investment with a Robo-Adviser (Continued)	
	(1)
Stock Value	0.0000232
	(0.58)
Bond Value	0.0000215
	(0.54)
ETF Value	0.0000236
	(0.59)
Mutual Fund Value	0.0000250
	(0.63)
Levered Prod. Value	0.000000340
	(0.01)
Derivative Value	0.0000290
	(0.73)
Buy Volume	0.00513***
	(4.72)
Sell Volume	-0.00663**
	(-2.91)
# Asset Classes	0.287***
	(12.71)
# Securities	-0.0116***
	(-4.44)
N	14,385
R^2	
Weighting Correction	Yes

t statistics in parentheses

Table 3.9: Summary Statistics of Active Robo-Adviser Users and their Nearest Neighbors: This table reports the Summary Statistics and the difference in Summary Statistics between the Robo-Adviser Users and their matched sample of the Control Group Retail Investors (Control).

Summary Statistics after Propensity Score Matching						
	Control		Robo-Adviser Users		Difference	
	mean	sd	mean	sd	b	t
Female (in %)	11.40	31.79	10.68	30.89	0.72	(1.26)
Male (in %)	61.49	48.67	62.85	48.32	-1.36	(-1.53)
Age	53.78	15.47	53.45	12.83	0.33	(1.28)
Years being Customer	0.86	4.76	0.90	4.83	-0.04	(-0.46)
Portfolio Value (in 1000)	81.37	108.64	81082.09	120.15	0.29	(0.14)
Portfolio Age	8.32	4.86	8.39	4.98	-0.07	(-0.79)
Stock Ind. (in %)	76.65	42.31	74.21	43.75	2.44**	(3.09)
Bond Ind. (in %)	16.56	37.18	16.71	37.31	-0.15	(-0.22)
ETF Ind. (in %)	47.51	49.94	49.76	50.00	-2.26*	(-2.46)
Mutual Fund Ind. (in %)	64.20	47.94	64.18	47.95	0.02	(0.02)
Levered Prod. Ind. (in %)	3.81	19.14	3.54	18.47	0.27	(0.78)
Derivative Ind. (in %)	15.65	36.34	15.94	36.60	-0.29	(-0.43)
Other Security Ind. (in %)	3.74	18.98	2.98	17.01	0.76*	(2.29)
Stock Value (in 1000)	36.98	77.71	36.87	79.15	0.11	(0.08)
Bond Value (in 1000)	3.50	14.70	3.62	14.91	-0.12	(-0.45)
ETF Value (in 1000)	14.48	34.56	13.47	30.36	1.01	(1.69)
Mutual Fund Value (in 1000)	24.46	45.38	25.09	46.89	-0.63	(-0.74)
Levered Prod. Value	107.11	715.76	111.98	747.07	-4.87	(-0.36)
Derivative Value (in 1000)	1.77	6.42	1.84	6.44	-0.07	(-0.57)
Other Security Value	64.77	529.47	78.07	590.42	-13.30	(-1.29)
# Buys	16.44	27.49	16.83	27.68	-0.39	(-0.76)
# Trades	21.21	36.42	21.78	36.23	-0.57	(-0.85)
# Sells	4.92	12.56	4.92	11.98	-0.00	(-0.00)
Buy Volume (in 1000)	38.33	98.18	39.14	92.92	-0.81	(-0.46)
Sell Volume (in 1000)	27.26	85.38	28.67	84.73	-1.41	(-0.91)
# Asset Classes	2.28	1.10	2.26	1.13	0.01	(0.64)
# Securities	10.86	10.48	10.61	10.88	0.25	(1.27)
N	5,936		5,936		11,872	

Table 3.10: Retail Investors Portfolio Characteristic Changes after Investing with the Robo-Adviser (on Propensity Score Matching): This table shows the regression results of a Difference-in-Difference regression. It compares the actively managed portfolio of the Robo-Adviser Users with the (actively managed) portfolio of Non-Robo-Adviser Users from 2016 (before the Robo-Advisor introduction) to 2018. It included only retail investors with an active portfolio in 2016. The dependent variables are: *Diversified Ind.* is an indicator being 1 if the retail investor holds ETFs or Mutual Funds in his (active) portfolio. *Asset Classes* counts the number of asset classes in the retail investor's portfolio. *ETF Share*, *MF Share* report the fraction of the active portfolio held in Mutual Funds or ETFs. *Trades*, *Buys*, *Sells* reports the number of trades/buys/sells during that year by the retail investor. The Weighting Correction refers to the Correction used to adjust for the Oversampling of the Robo-Advisor Users relative to the Non-Robo-Advisor Users. The Control group is constructed from the original sample via Propensity Score Matching and finding the nearest neighbor. The control variables include personal characteristics: Sex, Joint Account, Age and portfolio characteristics: Years as Customer, Years since opening the portfolio and the 2016 investment characteristics portfolio value, stock dummy, bond dummy, ETF dummy, Mutual Funds Dummy, Levered Product Dummy, Derivative Dummy, Buy Volume, Sell Volume, Number of Asset Classes, Number of Securities and Turnover. The standard errors are clustered at the individual level.

Change in Retail Investors Active Portfolio Characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Diversified Ind.	Asset Classes	ETF Share	MF Share	Trades	Buys	Sells
$D(Robo)$	-0.00665	-0.000579	-0.0177***	-0.00469	0.184	0.0172	0.00584
	(-1.23)	(-0.15)	(-3.55)	(-0.87)	(1.62)	(0.24)	(0.12)
$D(Robo \times \geq 2017)$	0.00462	0.0206	0.00543	0.0149***	1.201*	0.640	0.664***
	(0.87)	(1.34)	(1.47)	(4.18)	(2.42)	(1.73)	(3.43)
$D(Robo \times 2018)$	-0.00119	0.0313*	0.00369	-0.0121***	-1.510**	-0.831*	-0.549**
	(-0.25)	(2.56)	(1.24)	(-4.17)	(-2.82)	(-2.15)	(-2.75)
N	35,131	35,131	34,779	34,779	35,131	35,131	35,131
R^2	0.577	0.794	0.529	0.555	0.754	0.761	0.708
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighting Correction	No	No	No	No	No	No	No

t statistics in parentheses

Table 3.11: Retail Investors' Portfolio Characteristic Changes after Investing with the Robo-Adviser (with ZIP Code Fixed Effects): This table shows the regression results of a Difference-in-Difference regression. It compares the actively managed portfolio of the Robo-Adviser Users with the (actively managed) portfolio of Non-Robo-Adviser Users from 2016 (before the Robo-Advisor introduction) to 2018. It included only retail investors with an active portfolio in 2016. The dependent variables are: *Diversified Ind.* is an indicator being 1 if the retail investor holds ETFs or Mutual Funds in his (active) portfolio. *Asset Classes* counts the number of asset classes in the retail investor's portfolio. *ETF Share*, *MF Share* report the fraction of the active portfolio held in Mutual Funds or ETFs. *Trades*, *Buys*, *Sells* reports the number of trades/buys/sells during that year by the retail investor. The Weighting Correction refers to the Correction used to adjust for the Oversampling of the Robo-Advisor Users relative to the Non-Robo-Advisor Users. The control variables include personal characteristics: Sex, Joint Account, Age and portfolio characteristics: Years as Customer, Years since opening the portfolio and the 2016 investment characteristics portfolio value, stock dummy, bond dummy, ETF dummy, Mutual Funds Dummy, Levered Product Dummy, Derivative Dummy, Buy Volume, Sell Volume, Number of Asset Classes, Number of Securities and Turnover. The standard errors are clustered at the individual level. All specifications include *Postal Code* $\times$ *Year* Fixed Effects.

Change in Retail Investors Active Portfolio Characteristics							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Diversified Ind.	Asset Classes	ETF Share	MF Share	Trades	Buys	Sells
$D(Robo)$	-0.00500	0.0341***	0.000741	0.0187***	0.237	0.186	0.0466
	(-0.86)	(4.93)	(0.17)	(3.72)	(0.92)	(1.04)	(0.47)
$D(Robo \times \geq 2017)$	-0.00880	-0.0248*	0.00402	0.00719*	1.614***	0.910**	0.682***
	(-1.66)	(-1.97)	(1.30)	(2.35)	(3.87)	(3.07)	(3.98)
$D(Robo \times 2018)$	-0.00644	-0.00159	0.000210	-0.0108***	-1.664***	-1.218***	-0.410*
	(-1.43)	(-0.15)	(0.08)	(-4.13)	(-3.92)	(-4.04)	(-2.35)
N	42,138	42,138	41,777	41,777	42,138	42,138	42,138
R^2	0.764	0.834	0.694	0.744	0.787	0.795	0.751
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Postal Code $\times$ Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighting Correction	Yes	Yes	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

Bibliography

Abraham, Facundo, Sergio L. Schmukler, and Jose Tessada, "Robo-Advisors : Investing through Machines," *Research and Policy Briefs*, 2019, *1* (21), 1–4.

Afonso, Gara and Ricardo Lagos, "Trade Dynamics in the Market for Federal Funds," *Econometrica*, 2015, *83* (1), 263–313.

Ateş, Sinem, Ali Coşkun, M. Abdullah Şahin, and M. Levent Demircan, "Impact of Financial Literacy on the Behavioral Biases of Individual Stock Investors: Evidence from Borsa Istanbul," *Business and Economics Research Journal*, 2016, *7* (3), 1–19.

Badarinza, Cristian, John Y. Campbell, and Tarun Ramadorai, "International Comparative Household Finance," *Annual Review of Economics*, 2016, *8* (1), 111–144.

Bank of International Settlement, "Correspondence Banking," *Report by Committee on Payments and Market Infrastructures*, 2016.

_ , "Triennial Central Bank Survey of Foreign Exchange and Over-the-counter (OTC) Derivatives Markets in 2019," *Report*, 2019.

Barber, Brad M. and Terrance Odean, "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 2000, *55* (2), 773–806.

Barras, Laurent, Olivier Scaillet, and Russ Wermers, "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas," *Journal of Finance*, 2010, *65* (1), 179–216.

Bernhardt, Dan, Vladimir Dvoracek, Eric Hughson, and Ingrid M. Werner, "Why Do Larger Orders Receive Discounts on the London Stock Exchange?," *Review of Financial Studies*, 2004, *18* (4), 1343–1368.

Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer, "Is Unbiased Financial Advice to Retail Investors Sufficient? Answers from a Large Field Study," *Review of Financial Studies*, 2012, *25* (4), 975–1032.

Bjønnes, Geir H. and Neophytos Kathiziotis, "Price Discovery in Fragmented Electronic Markets: The Case of FX," *Working Paper*, 2016.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 2007, *115* (5), 707–747.

Chien, YiLi and Paul Morris, "Household Participation in Stock Market Varies Widely by State," *The Regional Economist*, 2017, *25* (3), 4–5.

Cocco, João F., Francisco J. Gomes, and Nuno C. Martins, "Lending Relationships in the Interbank Market," *Journal of Financial Intermediation*, 2009, *18* (1), 24 – 48.

Deutsches Aktien Insitut, "Aktionärszahlen des deutschen Aktieninstituts 2018," *Report*, 2019.

Deutsches Aktien Institut, "Mehr Aktionäre in Deutschland," *Report*, 2019.

D'Hondt, Catherine, Rudy De Winne, Eric Ghysels, and Steve Raymond, "Artificial Intelligence Alter Egos: Who Benefits from Robo-investing?," *Working Paper*, 2019.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, "The Value of Trading Relations in Turbulent Times," *Journal of Financial Economics*, 2017, *124* (2), 266 – 284.

_ , Francesco Franzoni, Amir Kermani, and Carlo Sommavilla, "The Relevance of Broker Networks for Information Diffusion in the Stock Market," *Journal of Financial Economics*, 2019, *134* (2), 419 – 446.

_ , Mark Egan, and Francesco Franzoni, "The Value of Intermediation in the Stock Market," *Working Paper*, 2019.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Pedersen, "Over-the-Counter Markets," *Econometrica*, 2005, *73* (6), 1815–1847.

D'Acunto, Francesco, Nagpurnanand Prabhala, and Alberto G Rossi, "The Promises and Pitfalls of Robo-Advising," *Review of Financial Studies*, 2019, *32* (5), 1983–2020.

Easley, David, Marcos Lopez de Prado, and Maureen O'Hara, "Discerning Information from Trade Data," *Journal of Financial Economics*, 2016, *120* (2), 269–285.

Edwards, Amy K., Lawrence E. Harris, and Michael S. Piwowar, "Corporate Bond Market Transaction Costs and Transparency," *Journal of Finance*, 2007, *62* (3), 1421–1451.

Euromoney, "Euromoney FX Survey 2019 Results Released: Press Release," *Euromoney*, 2019.

Evans, Martin D. D. and Dagfinn Rime, "Microstructure of Foreign Exchange Markets," *Working Paper*, 2019.

_ and Richard K. Lyons, "Order Flow and Exchange Rate Dynamics," *Journal of Political Economy*, 2002, *110* (1), 170–180.

Fama, Eugene F. and Kenneth R. French, "Luck versus Skill in the Cross-Section of Mutual Fund Returns," *Journal of Finance*, 2010, *65* (5), 1915–1947.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer, "The Emergence of Market Structure," *Working Paper*, 2017.

Financial Times, "Citi Plans to Withdraw from Two-Thirds of Foreign Exchange Platforms," *Financial Times*, October 2019.

Gale, William G., Benjamin H. L. Harris, and Ruth E. Levine, "Raising Household Saving: Does Financial Education Work?," *Social Security Bulletin*, 2012, *72* (2), 39–48.

Goldstein, Itay, Wei Jiang, and G Andrew Karolyi, "To FinTech and Beyond," *Review of Financial Studies*, 2019, *32* (5), 1647–1661.

Goldstein, Michael A. and Edith S. Hotchkiss, "Providing Liquidity in an Illiquid Market: Dealer Behavior in US Corporate Bonds," *Journal of Financial Economics*, 2020, *135* (1), 16 – 40.

Green, Richard C., Burton Hollifield, and Norman Schürhoff, "Dealer Intermediation and Price Behavior in the Aftermarket for new Bond Issues," *Journal of Financial Economics*, 2007, *86* (3), 643 – 682.

Grinblatt, Mark and Matti Keloharju, "The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set," *Journal of Financial Economics*, 2000, *55* (1), 43 – 67.

Grossman, Sanford J. and Merton H. Miller, "Liquidity and Market Structure," *Journal of Finance*, 1988, *43* (3), 617–633.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli, "Financial Advisors: A Case of Babysitters?," *Journal of Banking and Finance*, 2012, *36* (2), 509 – 524.

Hagströmer, Björn and Albert Menkveld, "Information Revelation in Decentralized Markets," *Journal of Finance*, 2019, *74* (6), 2751–2787.

Harris, Lawrence, "Transaction Costs, Trade Throughs, and Riskless Principal Trading in Corporate Bond Markets," *Working Paper*, 2015.

Hasbrouck, Joel and Richard M. Levich, "FX Liquidity and Market Metrics: New Results Using CLS Bank Settlement Data," *Working Paper*, 2019.

_ and _ , "Network Structure and Pricing in the FX Market," *Journal of Financial Economics*, 2021, *141* (2), 705–729.

Hendershott, Terrence and Anath Madhavan, "Click or Call? Auction versus Search in the Over-the-Counter Market," *Journal of Finance*, 2015, *70* (1), 419–447.

_ , Dan Li, Dmitry Livdan, and Norman Schürhoff, "Relationship Trading in Over-the-Counter Markets," *Journal of Finance*, 2020, *75* (2), 683–734.

_ , _ , _ , and _ , "True Cost of Immediacy," *Working Paper*, 2020.

_ , Dmitry Livdan, and Norman Schürhoff, "Are Institutions Informed about News?," *Journal of Financial Economics*, 2015, *117* (2), 249 – 287.

Hollifield, Burton, Artem Neklyudov, and Chester Spatt, "Bid-Ask Spreads, Trading Networks, and the Pricing of Securitizations," *Review of Financial Studies*, 2017, *30* (9), 3048–3085.

Jappelli, Tullio, "Economic Literacy: An International Comparison," *Economic Journal*, 2010, *120* (548), F429–F451.

King, Michael R., Carol L. Osler, and Dagfinn Rime, "The Market Microstructure Approach to Foreign Exchange: Looking back and Looking forward," *Journal of International Money and Finance*, 2013, *38*, 95 – 119.

Koestner, Maximilian, Benjamin Loos, Steffen Meyer, and Andreas Hackethal, "Do Individual Investors Learn from their Mistakes?," *Journal of Business Economics*, 2017, *87* (5), 669–703.

Kyle, Albert S., "Continuous Auctions and Insider Trading," *Econometrica*, 1985, *53* (6), 1315–1335.

Lee, Charles M. C. and Mark J. Ready, "Inferring Trade Direction from Intraday Data," *Journal of Finance*, 1991, *46* (2), 733–746.

Li, Dan and Norman Schürhoff, "Dealer Networks," *Journal of Finance*, 2019, *74* (1), 91–144.

Lusardi, Annamaria and Olivia S. Mitchell, "Financial Literacy around the World: An Overview," *Journal of Pension Economics and Finance*, 2011, *10* (4), 497–508.

_ and _ , "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 2014, *52* (1), 5–44.

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf, "Information Flows in Foreign Exchange Markets: Dissecting Customer Currency Trades," *Journal of Finance*, 2016, *71* (2), 601–634.

Moore, Michael, Andreas Schrimpf, and Vladyslav Sushko, "Downsized FX Markets: Causes and Implications," *BIS Quarterly Review*, 2016, *21* (4), 35–51.

Odean, Terrance, "Are Investors Reluctant to Realize Their Losses?," *Journal of Finance*, 1998, *53* (5), 1775–1798.

O'Hara, Maureen and Xing (Alex) Zhou, "The Electronic Evolution of Corporate Bond Dealers," *Journal of Financial Economics*, 2021, *140* (2), 368–390.

Oliver Wyman, "Entwicklungen von Robo Advice in Deutschland und Global," *Report*, 2019.

Ravikumar, B. and Evan Karson, "How Has Stock Ownership Trended in the Past Few Decades?," *St. Louis Fed on the Economy Blog*, 2018.

Riggs, Lynn, Esen Onur, David Reiffen, and Haoxiang Zhu, "Swap Trading after Dodd-Frank: Evidence from Index CDS," *Journal of Financial Economics*, 2020, *137* (3), 857 – 886.

Saar, Gideon, Jian Sun, Ron Yang, and Haoxiang Zhu, "From Market Making to Matchmaking: Does Bank Regulation Harm Market Liquidity?," *Working Paper*, 2019.

Schrimpf, Andreas and Vladyslav Sushko, "FX Trade Execution: Complex and Highly Fragmented," *BIS Quarterly Review*, 2019, *24* (4), 39–51.

Skiera, Vincent, "Centrality in OTC Markets, Liquidity Provision, and Prices," *Working Paper*, 2021.

_ , "Client-Dealer Intermediation in OTC Markets," *Working Paper*, 2021.

Upper, Christian and Andreas Worms, "Estimating Bilateral Exposures in the German Interbank Market: Is there a Danger of Contagion?," *European Economic Review*, 2004, *48* (4), 827 – 849.

Üslü, Semih, "Pricing and Liquidity in Decentralized Asset Markets," *Econometrica*, 2019, *87* (6), 2079–2140.

van Rooij, Maarten C.J., Annamaria Lusardi, and Rob J.M. Alessie, "Financial Literacy, Retirement Planning and Household Wealth," *Economic Journal*, 2012, *122* (560), 449–478.

Vogel, Sebastian, "When to Introduce Electronic Trading Platforms in Over-the-Counter Markets?," *Working Paper*, 2019.

Wang, Chaojun, "Core-Periphery Trading Networks," *Working Paper*, 2017.

Zitzewitz, Eric, "Paired Corporate Bond Trades," *Working Paper*, 2010.

Appendix A

Client-Dealer Intermediation in OTC Markets

A.1 Sample Construction for Data of Chapter 1

Table A.1 describes the sample used in the empirical analysis. There are 117,177 trades and 190,873 RFQ in the sample. The sample construction follows 4 steps. First, duplicate quotes were removed; according to my understanding, these arise when a dealer utilizes the last look option and cancels a quote before the trade is executed with another dealer. As can be seen, the utilization of last look, and thus quote duplication, happens very rarely (0.3% of trades) and does not affect the number of RFQs or trades. Second, RFQs with missing settlement information or non-standard settlement are removed. All trades are RFQs with non-standard settlement and not missing information. Since non-standard settlements are harder to trade and off-load with other clients or dealers, their transaction cost will not be representative of regular trades. Third, and similarly, trades outside of FX trading hours, which run from Monday mornings in Sydney to Friday evening in New York, are removed as most dealers will not be available for trading outside trading hours or will have difficulty hedging the trade. The biggest adjustment to the sample size comes through the availability of a midquote. Finally, I selected currency pairs for which I have midquotes, which are 178 currency pairs[1] for which midquote information is provided. These 178 currency pairs cover 96% of all trading in my data, and more than 95% of all trading in the FX market happens in these currency pairs[2] (Bank of International Settlement, 2019), which greatly expand upon the set of currency pairs studied in other studies

[1]Of these 190,873 RFQ 189,890 RFQ receive quotes and for 189,099 RFQ have a midquote. For trades there are 117,117 trades and 116,934 have a midquote. The full list of currency pairs are, in alphabetical order and ordering: AUDCAD, AUDCHF, AUDCNH, AUDCZK, AUDDKK, AUDEUR, AUDGBP, AUDHKD, AUDJPY, AUDMXN, AUDNOK, AUDNZD, AUDPLN, AUDSEK, AUDSGD, AUDTRY, AUDUSD, AUDZAR, CADCHF, CADCNH, CADCZK, CADDKK, CADEUR, CADGBP, CADHUF, CADJPY, CADMXN, CADNOK, CADNZD, CADPLN, CADSEK, CADSGD, CADTHB, CADTRY, CADUSD, CADZAR, CHFCNH, CHFCZK, CHFDKK, CHFEUR, CHFGBP, CHFHKD, CHFHUF, CHFILS, CHFJPY, CHFMXN, CHFNOK, CHFNZD, CHFPLN, CHFRON, CHFRUB, CHFSAR, CHFSEK, CHFSGD, CHFTHB, CHFTRY, CHFUSD, CHFZAR, CNHEUR, CNHGBP, CNHJPY, CNHNZD, CNHSEK, CNHSGD, CNHTRY, CNHUSD, CZKDKK, CZKEUR, CZKGBP, CZKHUF, CZKJPY, CZKNOK, CZKPLN, CZKRON, CZKSEK, CZKSGD, CZKUSD, CZKZAR, DKKEUR, DKKGBP, DKKJPY, DKKMXN, DKKNOK, DKKNZD, DKKPLN, DKKRUB, DKKSEK, DKKSGD, DKKTRY, DKKUSD, DKKZAR, EURGBP, EURHKD, EURHUF, EURILS, EURJPY, EURMXN, EURNOK, EURNZD, EURPLN, EURRUB, EURSEK, EURSGD, EURTHB, EURTRY, EURUSD, EURZAR, GBPJPY, GBPMXN, GBPNOK, GBPNZD, GBPPLN, GBPRON, GBPRUB, GBPSEK, GBPSGD, GBPTRY, GBPUSD, GBPZAR, HKDJPY, HKDSEK, HKDSGD, HKDUSD, HUFNOK, HUFPLN, HUFSEK, HUFSGD, HUFUSD, HUFZAR, ILSUSD, JPYMXN, JPYNOK, JPYNZD, JPYPLN, JPYRUB, JPYSEK, JPYSGD, JPYTRY, JPYUSD, JPYZAR, MXNNOK, MXNNZD, MXNPLN, MXNSEK, MXNSGD, MXNUSD, NOKNZD, NOKPLN, NOKSEK, NOKSGD, NOKTRY, NOKUSD, NOKZAR, NZDSEK, NZDSGD, NZDUSD, NZDZAR, PLNRUB, PLNSEK, PLNSGD, PLNTRY, PLNUSD, PLNZAR, RONUSD, RUBTRY, RUBUSD, SARUSD, SEKSGD, SEKTRY, SEKUSD, SEKZAR, SGDTHB, SGDUSD, SGDZAR, THBUSD, TRYUSD, TRYZAR, USDZAR. 97(=275-178) of these currency pairs also have the trade direction in the other direction.

[2]95% is derived from the OTC foreign exchange turnover by country and currency in April 2019, on a "net-gross" basis in local currency. The residual column is removed in the calculation and the sum is taken across all currency pairs in the table for which my data provides midquote information relative to the sum across all currency pairs in the table. Furthermore, the currency pairs for which I have no midquote information with any currency pair, thus no midquote can be calculated via triangle arbitrage, account for only 0.3% of the data.

in FX markets (Evans and Lyons, 2002; Menkhoff et al., 2016; Hagströmer and Menkveld, 2019; Hasbrouck and Levich, 2021). The data are winsorized at the 1- and 99-percentile of executed RFQs for both transaction cost and trade size.

A.2 Transaction Cost by Client Characteristics

Table A.2 shows how client size is related to the trading cost of clients that receive quotes from only one dealer. These clients choose to contact only one dealer, many maybe preferring an "even smaller" network. Table A.2 again confirms that smaller clients receive worse prices from a dealer. However, at the same time the estimated coefficient by client size is much larger than previously, now at -0.83 and -0.72 (with dealer fixed effects), compared to -0.52 and -0.25 (with dealer fixed effects) in Table 1.3 where all clients are included. These coefficients confirm that dealers provide better prices to clients with whom they trade more, as also shown in Table 1.5, Table 1.6 and Table 1.7, but also that in particular small clients, who are the clients that contact few and at times only one dealer, can receive much worse prices when trading little. Larger clients can increase their network and trade more with each dealer, leading to a more muted response of the client size and trading cost, but small clients need to contact at least one dealer, no matter how small they are, leading to potentially very large costs to them.

A.3 Procedure to Identify Dealers who also Request Quotes

To identify dealers that also request quotes from other dealers, thereby acting like clients, I make a few assumptions and rely on them trading in RPT some of the time. My dataset identifies each participant requesting a quote (requester), i.e., client or dealer, and provides each dealer a unique ID. However, the ID as a requester for a dealer who requests a quote need not match the ID of the dealer. I will refer to a dealer that requests a quote as a dealer being associated with a client ID. The process to identify dealers who also request quotes as clients comprises 5 steps:

1. **Dealers are Banks or Institutions**: I assume that a dealer, a market participant providing a quote, is either a bank or an institution, allowing me to restrict the set of possible clients to clients who are banks or institutional investors.

2. **Client and Dealer if identical have the same country**: If a dealer is associated with a client, then both of them are from the same country, so that I can further restrict the set of possible matches.

3. **A dealer does not trade with herself**: If a dealer is associated with a client then the dealer should not trade with herself; thus the dealer should not trade with the client ID of the associated client. This step allows me to remove any possible dealer-client pairs where the client requests a quote from the dealer. Focusing not just on RFQ, but extending the no trading requirement to Streaming, but also to the Forward and Swap markets on the platform.

4. **Identify the associated clients to dealers via RPT**: Without looking at the trade level, there are restrictions that can be made. In this step, I match for each client-dealer pair that still is possible all quotes provided by the dealer to all the quotes requested by the client, based on currency pair and trade direction, that match within a 5-second radius. In this step, I am trying to identify whether a dealer is associated to a client by looking at RPTs. I then hand check each pair, determining whether a dealer is associated to a client based on the number of identified RPTs, relative to all trading. I further look at the types of trades that are flagged as RPTs to avoid spurious connections, for example a very small trade in EURUSD being an RPT, but larger trades not having any activity close by.

5. **Check monotonicity of matches**: The order of IDs for clients and dealers are monotone, so if *Dealer* x is identified as *Client* y, then *Dealer* x^*, with $x^* > x$, will match to *Client* y^*, with $y^* > y$. I thus check whether the matching that I determined is also monotone, which it is.

These 5 steps allow me to identify 110 different dealers that also act as clients; most of them trade via RFQ (97), and some trade solely via Streaming (13). In total, 29% of all dealers in RFQ are identified as acting also as clients on this multi-bank platform.

A.4 Time Difference in Riskless Principal Transactions

Figure A.1 shows the difference in time stamps (in seconds) between the two trades in an RPT. A positive value means that the dealer's RFQ with another dealer is recorded after the client's RFQ with the dealer, whereas a negative number means that the client's RFQ is recorded after the dealer's RFQ with another dealer. The distributions of both matched quotes and matched trades are centered very closely around 0, with a majority being matched within 0.025 seconds, and very few matches taking place outside 0.5 seconds. The differences between trades and quotes in Figure A.1 also support that these matches are RPT. In cases where the matched quotes are trades, the dealer executes before the client, indicating that the dealer waits until she receives confirmation on her trade from another dealer, before confirming, and thus executing, the trade of the client. However, in cases where the matched quotes do not match, the dealer's RFQ is generally timed after the client's RFQ. In those cases, the dealer only cancels her RFQ after either 1) the client executes with another dealer and the dealer observes the completion of the RFQ, 2) the client cancels the RFQ and the dealer observes the cancellation, or 3) the client does not trade or cancel and instead waits until the termination of the RFQ. In all cases, the dealer only learns that the client no longer needs her quote when the client has completed or canceled her request. Thus, the dealer can only cancel her request after the client has done so. These results are confirmed when looking only at match makers in Figure A.2.

A.5 Appendix Chapter 1: Figures

Figure A.1: Difference in Trading Time between 2 legs of a Riskless Principal Transaction (RPT) The figure shows the time difference between the time stamps of the trade of the dealer with another dealer and the trade of the client with the dealer in an RPT. The green shows the distribution of matched quotes, while the blue shows the distribution of matched trades. Weighting is done by Trade Size.

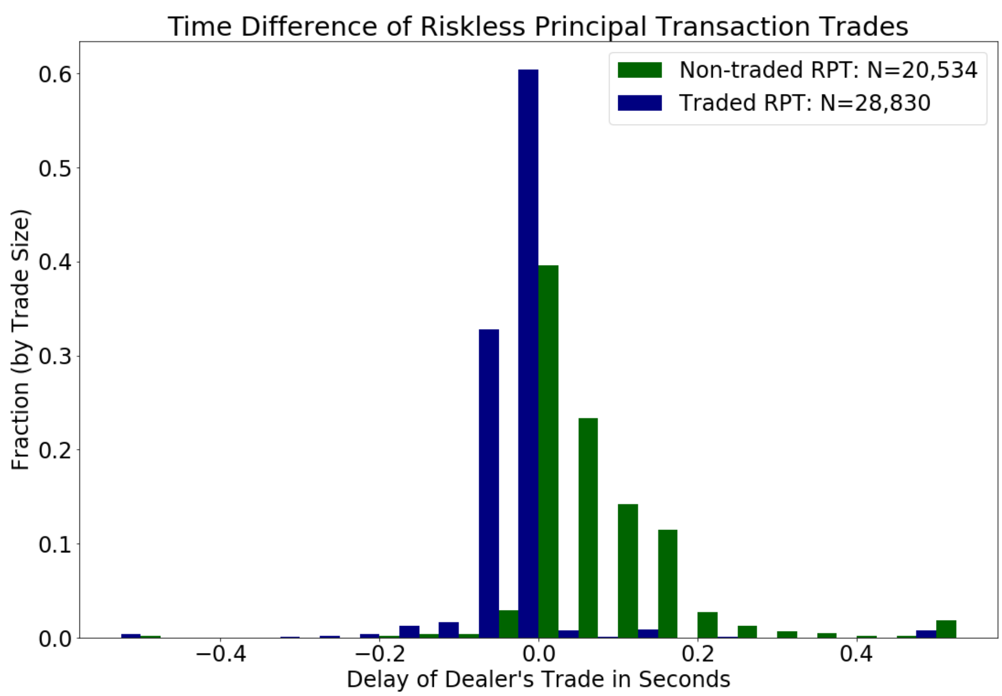
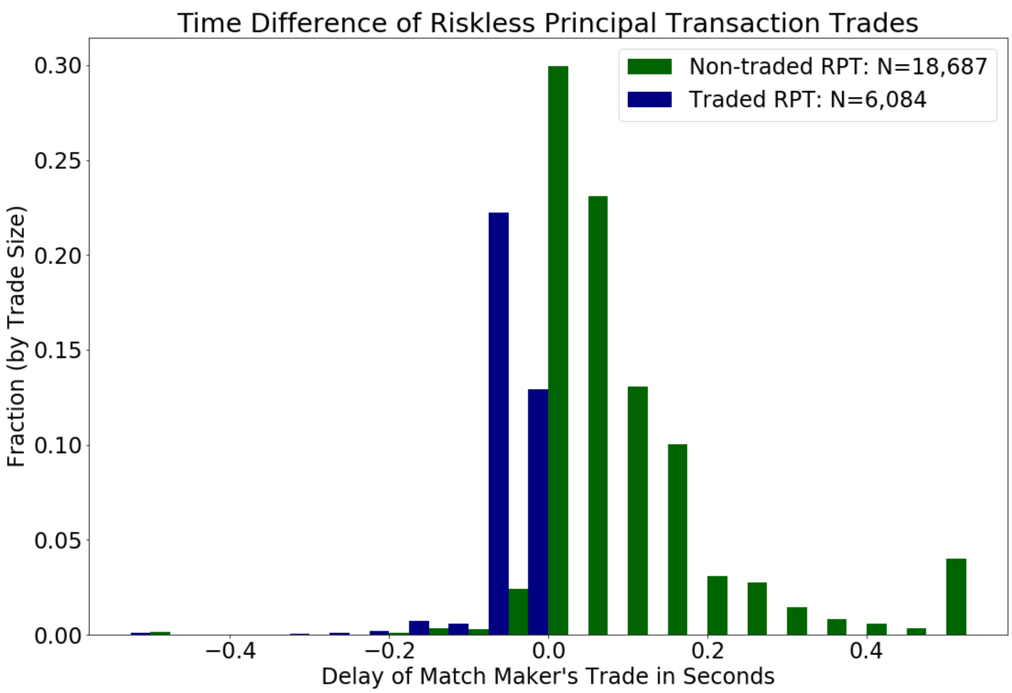


Figure A.2: Difference in Trading Time between 2 legs of a Riskless Principal Transaction (RPT) with a Match Maker The figure shows the time difference between the time stamps of the trade of the match maker with another dealer and the trade of the client with the match maker in an RPT. The green shows the distribution of matched quotes, while the blue shows the distribution of matched trades. Weighting is done by Trade Size.



A.6 Appendix Chapter 1: Tables

Table A.1: Sample Construction: This table shows the Sample that is used in the empirical analysis and how each manipulation of the sample changes the sample used in the empirical analysis.

Sample Construction			
	RFQ	Trades	Trades Volume
Raw data	206,854	131,331	98,571.22
Duplicated Quotes	4,061	357	399.39
Remaining Sample	206,854	131,331	98,171.82
Non-Standard or Missing Settlement	5,282	5,106	209.63
Remaining Sample	201,572	126,225	97,962.19
Outside Trading Hour (Weekend)	99	52	2.65
Remaining Sample	201,473	126,173	97,959.54
Currency Pair Restriction	10,600	8,996	4,088.19
(Pre-Winsorized) Final Sample	190,873	117,177	93,871.35
Final Sample	190,873	117,177	76,535.9

Table A.2: Transaction Cost for Clients Requesting one Quote: This table shows how client size is related to the transaction cost in an RFQ, focusing only on executed RFQs (trades) from clients requesting quotes from exactly one dealer. *Client Size* is the logarithm of a client's trade volume. *Trade Size in mil, Log Trade Size in mil* are the amount requested in mil EURO and its logarithm. *Institutional,Corporpartion* are dummy variables equal to 1 if the client is an institutional investor or corporation. Retail investors are subsumed by fixed effects, and banks are the unreported client type. *Buy & Sell Quote Provided* is a dummy variable equal to 1 if the RFQ requested quotes in both directions of the trade. All regressions include *Date, Time of Day* and *Currency Pair Fixed Effects*. Specification (2) introduces also *Dealer Fixed Effects*.

Trading Cost for Clients with at most One Dealer

	(1)	(2)
Client Size	-0.83***	-0.72***
	(0.01)	(0.01)
Trade Size in mil	0.14***	0.05***
	(0.02)	(0.01)
Log Trade Size in mil	-0.03***	-0.01
	(0.0)	(0.0)
Inst	-0.49***	-2.98***
	(0.07)	(0.09)
Corp	-0.3***	-1.74***
	(0.03)	(0.05)
Buy & Sell Quote Provided	0.18***	0.94***
	(0.06)	(0.06)
N	39,873	39,873
R^2	0.6326	0.821
Currency Pair F.E.	Yes	Yes
Date F.E.	Yes	Yes
Time of Day F.E.	Yes	Yes
Dealer F.E.	No	Yes

Table A.3: Simultaneous Quote Provision to Multiple Clients (based on Number of Dealers Contacted): This table shows the price difference between the quote a client contacting more dealers than a client contacting fewer dealers. The sample is constructed by looking at dealers that provide the same quote to two different clients for the same trade size and same currency pair within 1 second of each other. Row *All* looks at all occurrences where both the smaller network and larger network client trade with the dealer at some point in time, row *Larger Network Client trades more* conditions on the larger-network client trading more than the smaller-network client with the dealer providing the quote, while row *Larger Network Client trades less* conditions on both the smaller-network client trading more with the dealer. The size of the network is defined as the number of dealers from whom the client receives a quote. There are no cases in which the two clients are the same size or trade the same amount with the dealer, when converted to EUR. The columns are as follows: N is the number of observations used for each row, > *No Quote* is the fraction of observations where only the larger-network client does not receive a quote from the dealer, *<0* is the share of observations where the larger-network client receives a worse price, *=0* is the share of observations where the larger-network client and the smaller network client receive the same price, while *>0* is the share of observations where the smaller-network client receives a better price than the larger-network client (always provided by the same dealer). Finally, < *No Quote* is the fraction of observations where only the smaller network client receives no quote by the dealer, and ∀*No Quote* is the fraction of observations where both the smaller- and larger-network clients receive no quote from the dealer, again all conditional on both clients trading with the dealer at some point in time.

Simultaneous Quote Provision to Multiple Clients (for Client with more Dealers)							
	N	> No Quote	<0	=0	>0	< No Quote	∀ No Quote
All	245,887	2.1%	35.8%	30.1%	25.0%	4.3%	2.8%
Larger Network Client trades more	119,918	2.1%	20.6%	32.2%	35.8%	6.1%	3.3%
Larger Network Client trades less	125,969	2.1%	50.2%	28.1%	14.7%	2.5%	2.4%

Appendix B

Centrality in OTC Markets, Liquidity Provision, and Prices

B.1 Sample Construction for Data of Chapter 2

Table B.1 shows the sample construction of the sample used in this paper. Panel A shows the construction of the sample for RFQ orders, and Panel B shows the construction of the sample for Streaming orders. The raw data includes all trades at or above 100,000 EUR order size. Either sample restricts trading to within the trading hours, which run from Monday 8 am in Sydney (Sunday 10 pm UTC), Australia to Friday 4 pm in New York (Friday 9 pm UTC). The restriction removes 0.3% of trades in either sample (RFQ or Streaming sample).

Further, I restrict the sample to common currency pairs, which are currency pairs for which midquote data is available, and at least 50 trades occur per month[1]. This restriction leaves 184 currency pairs, all trade in RFQ, and 112 of these currency pairs also trade in Streaming. The restriction on common currency pairs removes 10.4% of trades in Streaming and 8.9% of executed orders in Streaming.

Besides these restrictions that apply to either sample, there is one more restriction specific to each sample, as the other sample always satisfies the restriction. Standard settlement in the FX spot market is spot, i.e., settlement is two days after the trade. All orders in Streaming have spot settlement, and 98.9% of RFQ trades do. Thus, to ensure comparability of transaction cost, the RFQ sample is restricted to spot-settled RFQ.

For Streaming, the client submits an order which executes against her personalized order book. This order need not be a market order, but can instead be a limit order, which is good until canceled[2]. To compare the transaction cost, I focus on orders that demand the same kind of immediacy as an RFQ order and the majority of Streaming orders. The restriction of immediacy for Streaming orders removes 3.1% of executed orders.

Some observations lack midquotes information. This lack of information affects 1.3% of RFQ trades and 8.6% of Streaming orders. I use all trades to calculate the centrality of a participant and statistics on liquidity provision. However, when it comes to the calculation of transaction cost, in the form of *Profit to Buyer*, I only use trades with a midquote to calculate the statistics, as only for those observations *Profit to Buyer* can be calculated. A subsample starting June 19th has much greater availability of midquote information, with less than 0.3% of RFQ trades and 2.3% of executed Streaming orders not having midquotes. Unreported results show that the results are unchanged for this subsample.

Finally, I winsorized the order size at 20mil EUR, which corresponds to the 99-percentile of the RFQ trade size distribution. The table reveals that this material impacts the trading volume executed via RFQ, reducing the trading volume by 12%. On the other hand, trading via Streaming is almost unaffected with a reduction of only 1% in both order size of executed volume and trade volume.

[1]The common currency pairs are:

[2]Most orders are immediate or cancel (IOC) limit orders, giving immediate, but at a bound worst price. So, despite demanding liquidity, these are not all market orders.

B.2 Matching of Liquidity Providers and Bank-Clients

The dataset identifies each participant requesting a quote (liquidity demander), i.e., client or dealer, and each participant providing a quote (liquidity provider) with a unique ID. However, the ID as a liquidity demander for a dealer who requests a quote need not match the ID of the dealer as a liquidity provider. I will refer to a dealer who requests a quote as a dealer being associated with a client ID. The process to identify dealers who also request quotes as clients comprises 5 steps:

1. **Dealers are Banks or Institutions**: I assume that a dealer, a market participant providing a quote, is either a bank or an institution, allowing me to restrict the set of possible clients to clients who are banks or institutional investors.

2. **Client and Dealer if identical have the same country**: If a dealer is associated with a client, then both of them are from the same country so that I can further restrict the set of possible matches.

3. **A dealer does not trade with herself**: If a dealer is associated with a liquidity demander, then the dealer should not trade with herself; thus, the dealer should not trade with the client ID of the associated liquidity demander. This step allows me to remove any possible liquidity provider-demander pairs where the liquidity demander requests a quote from the liquidity provider ID. Focusing not just on RFQ but extending the no trading requirement to Streaming and the Forward and Swap markets on the platform.

4. **Identify the associated liquidity demander to dealers via close-by trading**: Without looking at the trade level, some restrictions can be made. In this step, I check by hand each remaining possible liquidity demander ID and dealer ID pair. I want to assert beyond reasonable doubt that the liquidity demander is associated with the dealer ID. For this, I check whether the liquidity demander engages in the following behaviors: frequent riskless principal transactions (RPT) or hedging. In RPT, the dealer immediately passes on the trade to another dealer demanding liquidity, i.e., after the dealer ID provides liquidity to a liquidity demander, the liquidity demander ID immediately trades with another dealer in the same currency pair. Hedging includes more wide-ranging activities but is most commonly determined by the dealer ID trading with a client in a Forward and the liquidity demander ID immediately trading the same spot trade, so same trade direction and quantity. In a limited number of cases, the dealer hedges a large trade by slowly unloading the trade after trading with a liquidity demander, by herself demanding liquidity via the associated liquidity demander's ID.

5. **Check monotonicity of matches**: The order of IDs for clients and dealers are monotone, so if *Dealer* x is identified as *Client* y, then *Dealer* x^*, with $x^* > x$, will match to *Client*

y^*, with $y^* > y$. I thus check whether the matching that I determined is also monotone, which it is.

These 5 steps allow me to identify 140 different liquidity providers that also demand liquidity; almost all of them trade via RFQ (133), and almost half via Streaming (60). In total, I identify 36%($=\frac{140}{387}$) of all dealers as acting also as liquidity demanders on this multi-bank platform. The following describes which dealers are identified as liquidity providers as well.

B.2.1 Liquidity Providers Demanding Liquidity on the Platform

Figure B.1 shows the share of liquidity providers that also demand liquidity on the platform by the trade volume of the liquidity provider. The Top 25 liquidity providers are not matched to a liquidity demander ID on the platform, and outside the Top 25, roughly 35% of liquidity providers also demand liquidity. I believe that the largest dealers do not use the platform for trading. They run a single bank platform and have access to electronic limit order books and the interdealer market. If they want to trade small transactions, they can do so in the interdealer market, and large quantities are generally not traded on this platform. Therefore, it is unsurprising that I do not match them to a liquidity demander ID. The matching should instead be evaluated based on the liquidity demanding bank IDs, as the largest bank liquidity demander are likely smaller liquidity providers. The following subsection provides this comparison.

B.2.2 Liquidity Demand by Liquidity Providers on the Platform

The bank liquidity demanders are likely to provide liquidity to some of their clients. It should generally hold that the larger the bank liquidity demander, the more clients the bank (liquidity demander) has. Thus, the largest bank liquidity demander should also provide liquidity to their clients on the platform, and I should be more likely to match them to a liquidity provider ID. Table B.2 shows that the larger the bank liquidity demander, the more likely the bank liquidity demander also provides liquidity on the platform. Moreover, most larger bank liquidity demanders provide liquidity, with 70% of the 10 largest bank liquidity demanders also being a liquidity provider and around 50% of the 30 largest bank liquidity demanders. Within the next 70 largest bank liquidity demanders (remainder of Top 100), still, 40% are matched to a liquidity provider. Bank liquidity demanders that provide liquidity accounting for half of the liquidity demanded by banks (49%), showing that the matching is very good in matching large liquidity demanders.

The share of matched liquidity demanders drops unsurprisingly significantly for smaller banks (liquidity demanders) (outside the Top 200). These are either banks that rely very little on the platform (unlikely, see below) or are small and have few clients. With few clients and especially small clients, these banks will be less likely to have clients that access the platform, and the trade sizes will be so small that the bank is less likely to engage in the hedging of the position. Thus, the behavior to match these banks to a liquidity provider is more difficult to observe. Furthermore, there are many more banks than liquidity providers (514 vs. 387), showing that many banks are not liquidity providers (on this platform). Thus, despite the relatively fewer matches, the matching remains very good, as most of the smaller banks simply are not liquidity providers on the platform.

B.2.3 Liquidity Provision and Liquidity Demand by Liquidity Providers Demanding and Providing Liquidity

For liquidity providers that provide liquidity and supply liquidity on the platform, I can compare the liquidity that they provided versus the liquidity that they demanded. Figure B.3 shows exactly this joint distribution. Many liquidity providers are exactly on the 45°. These are match makers, as discussed in Skiera (2021b). However, of the 140 liquidity providers that also request liquidity, only 4 provide noticeably more liquidity than they request. The remaining 97% of liquidity providers demand at least as much liquidity as they provide. Thus, besides trading with clients on the platform, they likely have trading needs from hedging trades in Forward trades or trading with clients over the phone and exclusively use the platform to demand liquidity. Thus, providing concrete evidence that the vast majority of liquidity demanders likely exclusively use the platform for demanding liquidity.

B.3 Centrality, Liquidity Provision, and Transaction Cost in EURUSD

In this appendix, I repeat the primary analysis of the paper conducted in Section 2.4, Section 2.5, Section 2.6 and Section 2.8, but only looking at transactions in the currency pair EURUSD. Trades in EURUSD are the most common trades with 31.30% of trades (410,337 trades trading 533.22 bn) trading the currency pair EURUSD, involving 1,320 liquidity demanders and 306 liquidity providers. So two out of three participants trade in EURUSD, including 79% of all liquidity providers.

In the primary analysis, centrality groups are formed based on cumulative trading volume. To keep the groups comparable to the primary analysis, I create centrality groups based on each participant's trading volume in EURUSD and sort participants based on their EURUSD degree centrality, i.e., the degree centrality arising from using only trades in EURUSD. I follow Section 2.4 in the creation of centrality groups. Due to fewer participants trading in EURUSD, the groups contain slightly fewer participants. The most-central group contains 8 participants, and the next most-central group contains 25 participants. The third and fourth most-central groups are made up of the 34th to 80th and 81st to 200th most central participants, respectively. The three groups accounting for the last quintile of trading are made up of the 201st to 340th most-central participants and 341st to 600th most-central participants, with the most peripheral group having the 1,029 least-central participants (601+).

B.3.1 Centrality and Trade Prices in EURUSD

This subsection shows that a centrality premium emerges in trading when not conditioning on liquidity demand and liquidity provision. I.e., I show that the centrality premium is also found in specific currency pairs. For this Table B.2 reports the profit to a buyer conditional on the centrality of the buyer and seller in EURUSD trades. The rows show the centrality of the buyer, and the columns show the centrality of the seller. Table B.2 shows that the buyer makes a profit if the buyer is more-central than the seller. So while a Top 8 buyer makes a profit when trading with all sellers, a 34-80 buyer only makes a profit when trading with a buyer outside the two most central groups.

Just like in Table 2.5, in Table B.2 the profit a buyer makes is greater if the more central the buyer is relative to the seller. For example, a Top 8 buyer makes a profit of 0.07bps when trading with a 9-33 seller. This profit increases to 0.41bps when trading with the most peripheral sellers (601+). Finally, Table B.2 is roughly symmetric. The profits a participant from a more-central group makes when buying from a less-central group are similar to the losses the same less-central group makes when buying from the same more-central group.

The results thus confirm the results from Section 2.4.

B.3.2 Centrality and Liquidity Provision in EURUSD

This second subsection of Appendix B.3 studies who provides liquidity based on the centrality of the two participants in a trade. Table B.3 shows the fraction of trades in which the row-central participant demands liquidity from the column-central participants. The table shows the liquidity demand from the (weakly) less-central participant in the trade. It reveals that the less-central participants demand liquidity in more than 50% of trades in all but one case. Thus, generally, the more-central participant provides liquidity to the less-central participant.

Furthermore, the results reveal that the larger the difference in centrality the more-central participant provides liquidity relatively more often. For example, 601+ participants always demand liquidity from Top 80 participants, demand liquidity in more than 95% of trades from 81-200 and 201-340 participants, but only demand liquidity in 64% of trades from 341-600 participants.

Thus, Table B.3 confirms the results from Table 2.4, showing that the less-central participants tend to demand liquidity. In addition, unreported results show that outside the Top 33 participants, liquidity providers tend to demand more liquidity from other liquidity providers than they provide to liquidity demanders. The unreported results confirm the strong reliance of peripheral liquidity providers on interdealer trading in liquid currency pairs.

B.3.3 Centrality, Liquidity Provision, and Transaction Cost

In this subsection, I study how conditioning on the liquidity demand in EURUSD trades affects the relationship between centrality and transaction cost. To this end, Table B.4 reports the profit to a buyer based on the centrality of the buyer and the seller and conditioning on the participant that demanded liquidity. The prefix *Buyer* is used when the buyer demands liquidity. In those cases, the row represents the centrality of the buyer, while the column represents the centrality of the seller. When the prefix *Seller* is used, the seller demands liquidity. Then the row represents the centrality of the seller and the column the centrality of the buyer.

Table B.4 shows that demanding liquidity requires all participants to pay a fee. While sellers always pay a fee for demanding liquidity, buyers generally do. Only the most-central liquidity demanding buyers (9-33) may not pay a fee when trading with some liquidity providers[3]. As in Table 2.5, transaction costs are decreasing with the centrality of the liquidity demander. For example, for sellers demanding liquidity from 9-33 liquidity providers, a 9-33 seller pays transaction costs of 0.12bps, while 34-80 seller pays 0.15bps transaction cost. 81-200 and 201-

[3] 9-33 participants demand liquidity in only very few trades on this platform. Furthermore, compared to less-central liquidity providers 9-33 liquidity providers hold inventory and rely less on interdealer trading. Them making a profit when demanding liquidity may represent times when the liquidity provider has a significant inventory imbalance and is willing to trade at a preferable price relative to the jointly known midquote. Thus, 9-33 liquidity providers may be uniquely positioned to trade with other liquidity providers at these advantageous prices. Other liquidity providers may be willing to offer these advantageous prices, as the spread of information is very low, compared to the publically observable limit order book.

340 sellers pay 0.17bps and 0.19bps in transaction costs, respectively. Finally, 341-600 sellers pay 0.3bps in transaction costs, and the transaction costs for 601+ sellers are 0.51bps. The transaction costs are again mainly determined by the centrality of the liquidity demander and less determined by the centrality of the liquidity provider. For example, a liquidity demanding 9-33 seller pays between 0.06bps and 0.12bps transaction costs to all her liquidity providers, with 0.08bps paid to 3 very differently central liquidity provider groups.

For the transaction costs a liquidity demander faces based on the centrality of her liquidity provider, the results mirror the results in Table 2.5. For the 4 most central centrality groups, there are only slight differences in the transaction costs across liquidity providers and no pattern with respect to which central liquidity providers provide the lowest transaction costs. However, for the 2 most peripheral, i.e., least central, centrality groups, more central liquidity providers give lower transaction costs. For example, for a 601+ seller, the transaction costs when demanding liquidity from a Top 8 liquidity provider are 0.42bps. At the same time, they increase to 0.51bps to 9-33 liquidity providers and rise to 1.41bps when demanding liquidity from another 601+ liquidity provider. Thus, for peripheral liquidity demanders, a centrality discount is found.

Central liquidity demanders contact multiple liquidity providers when demanding liquidity. Especially, in a liquid currency pair, like EURUSD. The simultaneous contact of many dealers will diminish any differences in the average price provided by different liquidity providers, as the simulation in Section 2.7 showed. Those simulations were calibrated to match the statistics of EURUSD trades. Table B.5 thus shows the log-market share different liquidity providers have when providing liquidity to a participant, conditional on providing liquidity to the participant. The table thus tests whether a participant trades the same amount with all her liquidity providers. If this is not the case, it reports whether some central liquidity providers provide liquidity more often, i.e., they end up providing the best price more frequently.

Table B.5 shows that more central liquidity providers provide liquidity more often than less central liquidity providers. Conditional on a participant trading with two liquidity providers, the more-central liquidity provider trades more often with the participant. This finding results from the decreasing coefficients on centrality groups. A Top 8 liquidity providers trade $49\%(= e^{0.4} - 1)$ more than an average liquidity provider to a participant, and 9-33 liquidity providers still trade $8\% = (e^{0.08} - 1)$ more with the participant than the average liquidity provider. However, peripheral liquidity providers, i.e., all remaining liquidity providers, trade less than the average liquidity provider. They trade at least $13\%(= e^{-0.14} - 1)$ less than the average liquidity provider and up to $24\%(= e^{-0.28} - 1)$ less than the average liquidity provider for 341-600 liquidity providers.

Taken together, Table B.4 and Table B.5 show that more central liquidity providers provide better prices on average, in line with Table 2.5 and Table 2.8. The results thus show that conditional on demanding liquidity, more central liquidity providers provide better quotes on average. Thus, Appendix B.3 confirms the finding of the primary analysis and shows that the findings hold in individual currency pairs. Appendix B.3 shows that a liquidity discount/premium is observed in EURUSD trades when conditioning/not conditioning on liquidity demand.

B.4 Appendix Chapter 2: Figures

Figure B.1: Share of Liquidity Providers Demanding Liquidity: This figure shows the (equal-weighted) share of liquidity providers (dealers) that not only provide liquidity but also demand liquidity on the platform by the size of the dealer. The x-axis shows the ranks of liquidity providers in each group. The brackets report the share of liquidity that liquidity providers in each group provide. Liquidity providers are ranked based on the amount of provided liquidity in EUR.

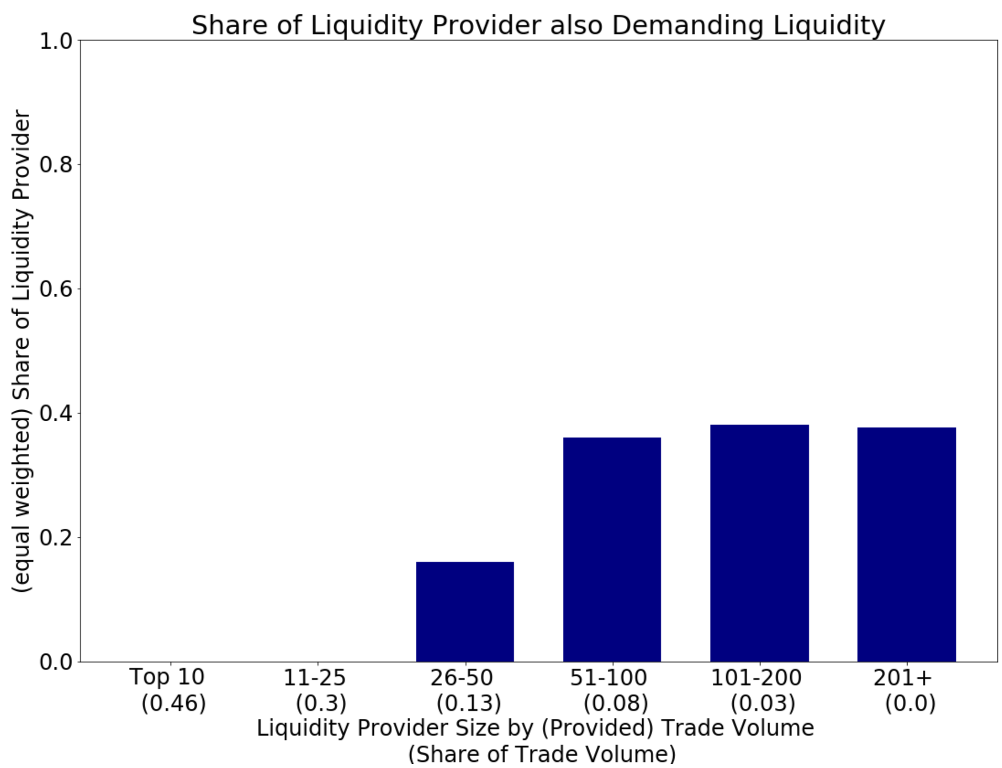


Figure B.2: Share of (Liquidity Demanding) Banks that Provide Liquidity: This figure shows the (equal-weighted) share of (liquidity demanding) banks that not only demand liquidity but also provide liquidity on the platform and are thus a (bank-)dealer. The x-axis shows the ranks of banks in each group and in brackets the share of trade volume, where banks demand liquidity, by banks in that group. Banks are ranked based on the amount of trade volume where a bank demands liquidity.

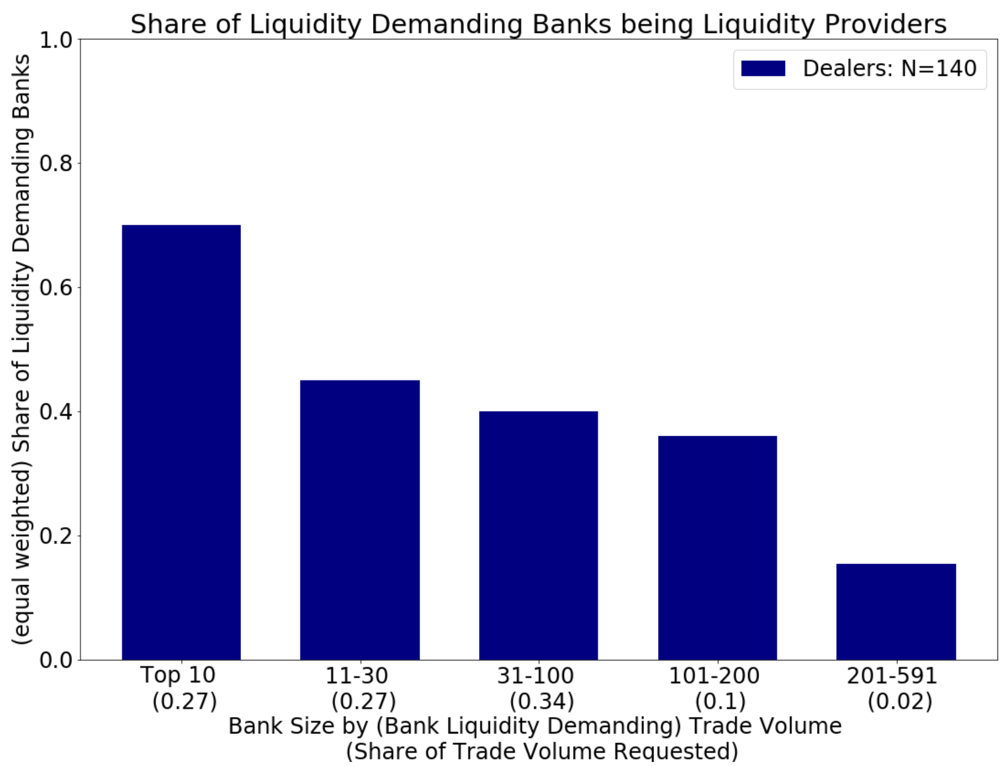
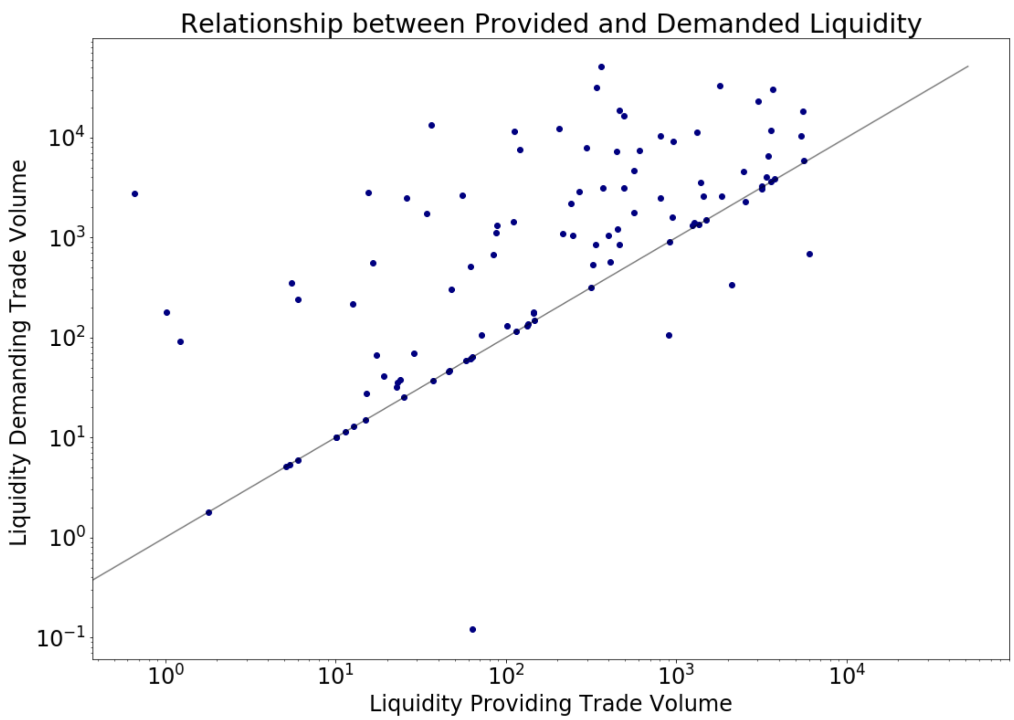


Figure B.3: Joint Distribution of Liquidity Providing and Liquidity Demanding Dealers: This figure shows the joint distribution of the amount of liquidity that is requested and the amount of liquidity that is provided by banks that request and provide liquidity.



B.5 Appendix Chapter 2: Tables

Table B.1: Sample Construction: This table shows the sample construction used in this research. The raw data includes all trades at or above 100k EURUSD. **Panel A:** shows the sample construction for RFQ Orders, while **Panel B:** shows the sample construction for Streaming Orders. **Panel A**: reports the number of RFQ (*No. RFQ*), the number of RFQ that lead to trades *No. Trades*), and the trade volume in mil EUR of the RFQ that lead to trades (*Trade Volume*). Restrictions for the sample construction are: Trades need to be within the trading hour, settlement needs to be standard (Spot settlement), and restricted to common currency pairs. **Panel B**: reports the number of orders in Streaming (*No. Orders*), the number of Orders that lead to at least one trade (*No. Executed*), the number of trades by orders that lead to at least one trade (*No. Trades*), the quantity that the liquidity demander wanted to trade in her order for orders that lead to at least one trade (*Executed Size*) and the trade volume in mil EUR of Orders that lead to a trade (*Trade Volume*). Restrictions on the sample construction are: Trades need to be within the trading hour, the submitted order needs to execute or cancel immediately, and restricted to common currency pairs. In both samples, I winsorized the order size at 20mil.

Panel A: Sample RFQ Orders

	No. RFQ	No. Trades	Trade Volume
Raw data	712,993	456,044	752,742.0
Outside Trading Hours	3,150	1,473	1,138.0
Remaining Sample	709,843	454,571	751,604.0
Non-Standard or Missing Settlement	7,411	4,953	3,563.0
Remaining Sample	705,746	451,075	750,219.0
Currency Pair Restriction	67,930	47,010	61,185.0
Final Sample	637,816	404,065	689,034.0
Midquote Missing	24,246	5,264	5,561.0
Final Sample (Winsorized)	637,816	404,065	606,382.0
From June 19th	494,043	319,137	551,935.0
From June 19th Midquote Missing	4,619	859	653.0
From June 19th (Winsorized)	494,043	319,137	482,581.0

Panel B: Sample Streaming Orders

	No. Orders	No. Executed	No. Trades	Executed Size	Trade Volume
Raw data	1,093,696	1,055,295	1,161,865	1,161,949.0	1,141,626.0
Outside Trading Hours	4,084	3,893	4,463	4,673.0	4,220.0
Remaining Sample	1,089,612	1,051,402	1,157,402	1,157,276.0	1,137,406.0
Immediacy Restriction	33,242	33,096	65,952	87,232.0	79,898.0
Remaining Sample	1,056,370	1,018,306	1,091,450	1,070,044.0	1,057,507.0
Currency Pair Restriction	94,329	90,704	94,919	71,305.0	70,939.0
Final Sample	962,041	927,602	996,531	998,739.0	986,569.0
Midquote Missing	125,737	80,203	88,294	99,802.0	96,116.0
Final Sample (Winsorized)	962,041	927,602	996,531	987,857.0	976,687.0
From June 19th	751,862	725,103	776,699	766,093.0	756,758.0
From June 19th Midquote Missing	51,757	16,511	17,944	21,727.0	19,498.0
From June 19th (Winsorized)	751,862	725,103	776,699	758,583.0	750,005.0

Table B.2: Profit to Buyer by Centrality in EURUSD: This table looks at the profit a certain central buyer (row) makes when trading with a certain central seller (column). Profit is measured as the difference in log basis points, between the price at which the buyer buys the currency pair and the midquote for the currency pair, i.e. the half-spread paid by the buyer. The values are the regression coefficient where the profit of a certain buyer is regressed on buyer centrality x seller centrality dummy variables. The regression includes 410,337 observations and further includes Date and Time of Day fixed effects, Currency Pair fixed effects, as well as trade mechanism fixed effects. The sample is constructed looking at all RFQ and Streaming trades with an trade size of at least 100,000 EUR. For Streaming trades, the order size is taken as the trade size, so the total quantity the initiator wants to trade, not just the particular amount traded with the dealer in that trade.

Profit to Buyer by Centrality in EURUSD

	Top 8	9-33	34-80	81-200	201-340	341-600	601+
Top 8		0.07***	0.13***	0.12***	0.15***	0.24***	0.41***
		(0.01)	(0.0)	(0.0)	(0.0)	(0.01)	(0.02)
9-33	0.03**	0.06***	0.15***	0.16***	0.18***	0.29***	0.5***
	(0.01)	(0.02)	(0.0)	(0.0)	(0.0)	(0.01)	(0.02)
34-80	-0.02***	-0.05***	0.06***	0.08***	0.09***	0.26***	1.34***
	(0.0)	(0.01)	(0.0)	(0.0)	(0.01)	(0.02)	(0.02)
81-200	-0.01**	-0.05***	0.0	0.03***	0.07**	0.15***	0.77***
	(0.0)	(0.0)	(0.0)	(0.01)	(0.03)	(0.03)	(0.02)
201-340	-0.03***	-0.07***	0.03***	-0.05**	0.11***	0.3***	1.39***
	(0.0)	(0.0)	(0.0)	(0.02)	(0.02)	(0.03)	(0.05)
341-600	-0.11***	-0.13***	-0.11	-0.01	-0.2***	0.1***	1.18***
	(0.01)	(0.01)	(0.07)	(0.03)	(0.04)	(0.03)	(0.08)
601+	-0.28***	-0.29***	-0.99***	-0.92***	-1.6***	-0.74***	0.02
	(0.01)	(0.02)	(0.02)	(0.01)	(0.05)	(0.06)	(0.04)

Table B.3: Liquidity Demand between Centrality Pairs in EURUSD: This symmetric table shows the share of trades between two participants with centrality corresponding to a given *row* and *column*, in which the *row*-centrality participant demands liquidity from the *column*-centrality participant.

Liquidity Demand between Centrality Pairs in EURUSD

	Top 8	9-33	34-80	81-200	201-340	341-600	601+
Top 8							
9-33	1.0	0.5					
34-80	1.0	1.0	0.5				
81-200	1.0	1.0	0.93	0.5			
201-340	1.0	1.0	0.55	0.4	0.5		
341-600	1.0	0.97	0.65	0.5	0.62	0.5	
601+	1.0	1.0	1.0	0.99	0.95	0.64	0.5

Table B.4: Profit to Buyer by Centrality Conditional on Liquidity Demand in EURUSD: This table looks at the profit a buyer with a certain centrality makes when trading with a seller with a certain centrality. The rows indicate the liquidity-requesting participant, i.e., the trade initiator. In cases with the prefix *Buyer*, the row represents the buyer in the transaction, and the column represents the seller in the transaction. In cases with the prefix *Seller*, these are switched, with the row representing the seller and the column representing the buyer. However, unchanged is that each cell represents the profit to the buyer. Profit is measured as the difference in log basis points between the price at which the buyer buys the currency pair and the midquote for the currency pair, i.e., the half-spread paid by the buyer. The reported values are the regression coefficient where I regress the profit of a certain buyer on buyer centrality x seller centrality dummy variables, conditional on the initiation of the trade. The regression includes 410,337 observations and further includes Date and Time of Day fixed effects, Currency Pair fixed effects, trade mechanism fixed effects, and trade size controls (linear and log of trade size). The coefficients on trade size (in mil) are: Linear -0.0 (0.0) and Log 0.01 (0.0). The sample is constructed looking at all RFQ and Streaming trades with a trade size of at least 100,000 EUR. For Streaming trades, the trade size is the order size, not just the amount traded with the dealer in that trade. The standard errors are reported below the regression coefficient. Standard errors are clustered at the *Currency Pair* $\times$ *Time of Date* level.

Profit to Buyer by Centrality Conditional on Liquidity Demand in EURUSD

	Top 8	9-33	34-80	81-200	201-340	341-600	601+
Buyer: Top 8							
Buyer: 9-33	0.03**	-0.05*	0.03	-0.01	0.07***	0.04***	
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	
Buyer: 34-80	-0.01***	-0.05***	-0.01**	-0.03***	-0.06***	-0.06**	-0.1
	(0.0)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.08)
Buyer: 81-200	-0.0	-0.04***	-0.01	-0.09***	-0.13***	-0.09***	-0.21***
	(0.0)	(0.0)	(0.0)	(0.01)	(0.02)	(0.02)	(0.06)
Buyer: 201-340	-0.02***	-0.06***	-0.03***	-0.29***	-0.07***	-0.07***	-0.2**
	(0.0)	(0.0)	(0.01)	(0.05)	(0.02)	(0.02)	(0.1)
Buyer: 341-600	-0.1***	-0.13***	-0.4***	-0.24***	-0.45***	-0.16***	-0.66***
	(0.01)	(0.01)	(0.04)	(0.03)	(0.08)	(0.04)	(0.04)
Buyer: 601+	-0.27***	-0.28***	-0.99***	-0.92***	-1.63***	-1.47***	-1.34***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.05)	(0.04)	(0.04)
Seller: Top 8							
Seller: 9-33	0.08***	0.12***	0.08***	0.08***	0.09***	0.06***	
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	
Seller: 34-80	0.14***	0.15***	0.13***	0.13***	0.1***	0.25*	0.19***
	(0.0)	(0.0)	(0.0)	(0.01)	(0.0)	(0.14)	(0.02)
Seller: 81-200	0.12***	0.17***	0.1***	0.14***	0.1***	0.22***	0.43**
	(0.0)	(0.0)	(0.0)	(0.01)	(0.01)	(0.03)	(0.17)
Seller: 201-340	0.16***	0.19***	0.17***	0.33***	0.18***	0.09***	0.28*
	(0.0)	(0.0)	(0.01)	(0.04)	(0.02)	(0.01)	(0.16)
Seller: 341-600	0.25***	0.3***	0.37***	0.4***	0.46***	0.33***	1.03***
	(0.01)	(0.01)	(0.03)	(0.06)	(0.04)	(0.03)	(0.1)
Seller: 601+	0.42***	0.51***	1.36***	0.79***	1.53***	2.7***	1.41***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.07)	(0.05)

Table B.5: Log Market Share by Liquidity Provider in EURUSD: This table looks at the market share of a liquidity provider centrality group. The table reports the regression coefficients when the log market share of a liquidity provider to a liquidity demander is regressed on the liquidity demander's log number of liquidity providers and the centrality of the liquidity provider. I calculate market share as the share of trades a liquidity provider executes with a liquidity demander as a fraction of all the liquidity demander's trades. The sample of trades is constructed by looking at Streaming and RFQ trades with a trade size of at least 100,000 EUR and trading the EURUSD currency pair. For Streaming trades, the trade size is the order size, not just the amount traded with the dealer in that trade.

Log Market Share by Liquidity Provider in EURUSD	
	(1)
Log Liquidity Provider	-1.29***
	(0.01)
Top 8	0.4***
	(0.04)
9-33	0.08**
	(0.04)
34-80	-0.14***
	(0.05)
81-200	-0.22***
	(0.06)
201-340	-0.16**
	(0.08)
341-600	-0.28***
	(0.1)
601+	-0.18*
	(0.1)
N	6,099.0
R^2	0.5633

Appendix C

Robo-Advisers: Household Stock Market Participation and Investment Behavior

C.1 Sample Construction for Data of Chapter 3

I outline how I arrive from the data provided to me to my final sample (Table C.1). My sample consists of 9,767 control group retail investors and 9,551 retail investors using the Robo-Adviser (Robo-Adviser users). First, I remove 27 Robo-Adviser Users due to inconsistencies in the portfolio age for the Robo-Adviser's portfolio between 2017 and 2018. This removal leaves me with 10280 observations Robo-Adviser Users.

Next, I remove Robo-Adviser Users that have opened an account with the Robo-Adviser at the end of 2017 but are not using the account. The cut-off is an investment below 7000 Euro, but mainly contains investors with no investment in the Robo-Adviser. This restriction removes 292 observations.

Further, to make statements about the investment behavior pre-Robo-Adviser introduction, I need to observe the investors in 2016. I, therefore, restrict myself to investors that are bank customers in 2016. This restriction removes 770 observations and leaves me with a final sample of 9,767 Control retail investors and 9,551 Robo-Adviser users.

All data is winsorized at the 1% and 99% every year. The value of the active portfolio is not winsorized and is instead replaced by the sum of the winsorized value of portfolio positions, i.e. $\textit{Portfolio Value}_{i,t} = \textit{winsorized Stock Value}_{i,t} + \textit{winsorized Bond Value}_{i,t} + \cdots$. The portfolio values are therefore changed at the right-tail.

C.2 Seasonal Variation in the Number of New Participants in Financial Markets

Figure C.1: Distribution of New Participants in Financial Markets across the Year: This figure depicts the fraction of new financial markets participants in years 2016 and 2017 from the control group that opened a portfolio in a particular month. By the data construction there are no retail investors in the control group opening an account in December 2017.

C.3 Appendix Chapter 3: Tables

Table C.1: Sample Construction: This table shows how the Used Sample is constructed from the provided data. One observation here is one retail investor, observed potentially multiple times.

Sample Construction		
	Control	Robo-Adviser Users
Provided Sample	10,000	10,307
Removing Robo-User Inconsistencies	10,000	10,280
Removing non-Robo using Retail Investors	10,000	9,988
Removing new Bank Customers in 2017	9,767	9,551
Sample of active Retail Investors in 2016	9,032	6,089
Sample of Non-active Retail Investors in 2016	968	3,899