

Centralizing Over-the-Counter Markets?

by Jason Allen¹ and Milena Wittwer²

¹ Economic and Financial Research Department
Bank of Canada, Ottawa, Ontario K1A 0G9

² Stanford University

jallen@bankofcanada.ca, wittwer@stanford.edu



Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

ISSN 1701-9397

©2021 Bank of Canada

Acknowledgements

The views presented are those of the authors and not necessarily those of the Bank of Canada. Any errors are our own. We thank experts at the Bank of Canada—in particular, Corey Garriott and the Triton project team for critical insights into debt management—and at CanDeal for providing detailed institutional information. We also thank Darrell Duffie, Liran Einav and Matthew Gentzkow in addition to Luis Armona, Markus Baldauf, Timothy Bresnahan, Nina Buchmann, Jose Ignacio Cuesta, Piotr Dworczak, Jakub Kastl, David K. Levine, Negar Matorian, Paul Milgrom, Monika Piazzesi, Paulo Somaini, Robert Wilson, Ali Yurukoglu, and all participants in the IO, finance, and macro student workshops at Stanford.

Abstract

In traditional over-the-counter (OTC) markets, investors trade bilaterally through intermediaries referred to as dealers. An important regulatory question is whether to centralize OTC markets by shifting trades onto centralized platforms. We address this question in the context of the liquid Canadian government bond market. We document that dealers charge markups even in this market and show that there is a price gap between large investors who have access to a centralized platform and small investors who do not. We specify a model to quantify how much of this price gap is due to platform access and assess welfare effects. The model predicts that not all investors would use the platform even if platform access were universal. Nevertheless, the price gap would close by 32%–47%. Welfare would increase by 9%–30% because more trades are conducted by dealers who have high values to trade.

Topics: Financial institutions; Market structure and pricing

JEL codes: D40, D47, G10, G20, L10

1 Introduction

Each year, trillions of dollars’ worth of bonds, mortgage-backed securities, currencies, commodities, and derivatives are traded in over-the-counter (OTC) markets. Unlike centralized markets, such as stock exchanges, OTC markets are considered to be decentralized because buyers must search for sellers one by one in order to trade. Most OTC markets, therefore, rely on large financial institutions (dealers) to intermediate between investors (such as firms, banks, public entities, or individuals).

A series of antitrust lawsuits that accused dealers of abusing market power when trading with investors, combined with dramatic events during the COVID-19 crisis, raised questions regarding whether and how to centralize OTC markets.¹ A popular proposal is to shift trading onto multi-dealer platforms, on which investors run auctions with dealers. Yet even though this approach has already been adopted in some markets, it is unclear whether it has sizable effects on prices and welfare.² This is because although the platform can foster competition between dealers, it might be costly to use, especially if trading on the platform is not anonymous.

We study dealer market power and evaluate price and welfare effects from centralizing OTC markets with trade-level data on the Canadian government bond market. This market is considered to be close to efficient because it is highly liquid and features low price uncertainty. Nevertheless, we document sizable markups. We also show that large (institutional) investors, who have access to a platform, pay systematically lower prices than small (retail) investors, who do not. Our main analysis quantifies the role of platform access in driving the price gap and the changes in market outcomes and welfare that could result if platform access were universal.

Three features render the Canadian government bond market a particularly attractive setting for our research question. First, as a government bond market, it

¹Logan (2020) discusses the events in the U.S. Treasury market during the COVID-19 crisis and possible market reforms. For an overview of antitrust litigation, see “The Manipulation Monitor” at www.schlamstone.com, accessed on 05/06/2021.

²For example, the Dodd-Frank Act mandates that standardized derivatives must be traded on platforms called swap execution facilities (SEFs). Another example is the European Union’s Markets in Financial Instruments Directive II (MiFID II).

lies at the core of the financial system. Second, a multi-dealer platform—which is like platforms in many other OTC markets, including the largest ones in the United States and Europe—exists, but not all investors have access to it. This is a useful institutional feature we exploit in our empirical strategy. Third, a reporting regulation allows us to observe trade-level data, so that we can zoom in on each individual trade, unlike prior studies on government bond markets.³

The data cover essentially all trades that involve Canadian government bonds, as well as bidding data from all primary auctions in which the government issues bonds. The data set is unique in that it includes identifiers for market participants and securities, so that we can trace both through the market. We observe the time, price, and size of trades and know whether a trade was executed bilaterally or on the platform and whether an investor is retail or institutional. In addition, we collect bid and ask quotes that are posted on Bloomberg and therefore publicly available. They indicate at what prices investors can trade and serve as market values of each bond.

Our trade-level data allow us to document novel facts. We show that dealers charge markups over market value. These markups vary across investors and are systematically smaller for institutional than for retail investors. To test whether this is because of platform access, we use an event study design to show that the prices of an investor who (exogenously) loses platform access drop by an amount that is eight times the bid-ask spread. This raises the possibility that making platform access universal could lead to better prices. It neglects the fact, however, that dealers might respond by adjusting platform quotes.

To assess all price effects and quantify welfare gains when centralizing the market, we introduce a model in which dealers and investors have different values for realizing trade. They play a two-stage game. First, dealers simultaneously post (indicative) quotes at which they are willing to trade on the platform. Then, institutional investors

³Relative to other OTC markets, we know little about government bond markets because trade-level data are not readily available. The U.S., for example, began collecting trade-level data in mid-2017 but does not make it accessible to academics. Some countries granted access to data on parts of their Treasury market (Dunne et al. (2015); Monias et al. (2017); Kondor and Pintér (2019); De Roure et al. (2020)).

can enter the platform to run an auction among dealers. They expect to trade at the posted quotes but have to pay a cost—which, for instance, reflects concerns about revealing information in the auction. This can be costly, because it might harm future prices or an existing relationship with the dealer. Alternatively, institutional investors trade bilaterally at a price equal to their value—an assumption that is relaxed in model extensions. Retail investors can only trade bilaterally.

In estimating the model, we face the common challenge that prices (here quotes) are endogenous. Our solution is to construct a new cost-shifter instrument that changes the dealer’s costs to sell but not the investor demand. For this, we use bidding data on primary auctions, in which dealers buy bonds from the government to sell them at a higher price to investors. When a dealer wins more than she expected to win when bidding, she can more cheaply satisfy investor demand, either because of how others bid in the auction or because the government issued more than the dealer expected. Thus, how much more the dealer wins relative to what she expected to win represents an exogenous cost-shifter, which we construct with estimation techniques from the multi-unit auctions literature.

With the model, we can study what happens to prices and welfare if we allow all investors to trade on the platform via a counterfactual. Loosely speaking, institutional investors who can already trade on the platform act as a control group, and retail investors are the treatment group. More broadly, this counterfactual informs us about what happens when opening a platform or removing entry barriers.

Our main findings are twofold. First, we find that universal platform access reduces the price gap between retail and institutional investors by at least 30%. The gap does not close entirely, because some retail investors (40%) choose not to use the poorly designed platform. Platform competition is imperfect and, more importantly, it is costly to trade on the platform. This is in line with the concern raised by industry experts that investors are deterred from using platforms because they are reluctant to reveal their names.⁴

⁴Dealers are accused of “long [having] arranged trades bilaterally with investors away from platforms” (*Financial Times* (2015)). One worry is that “the loss of anonymity deters access to platforms in practice” (Managed Funds Association

Second, we find substantial welfare gains from providing universal access to the platform: Total gains from trade, our measure of welfare, increase by 9%–30%, or 0.8–2.6 basis points (bps) of GDP. The gains from trade increase because market participants have heterogeneous values for realizing trade, even though there is little uncertainty about the market value of the bond. Since the platform allows investors to access all dealers, more trades are intermediated by dealers who have high values to trade—for instance, because they seek to offload their inventory positions.

Taken together, our results have valuable policy implications for OTC markets. They emphasize that granting platform access alone does not shift all bilateral trades onto the platform, due to platform usage costs. A possible solution put forward by industry experts is to allow investors to trade anonymously on the platform. This could reduce privacy concerns and therefore the cost of using the platform. Moreover, our findings highlight that there are frictions that prevent dealers from buying and selling. We have seen this in the recent COVID-19 crisis, when dealers failed to absorb the excess supply of U.S. Treasury bonds on their balance sheets. These frictions can be reduced when shifting trades onto centralized platforms. We expect this to be true for many other OTC markets for standardized financial products (such as simple interest rate swaps or credit derivative index products), in which welfare gains are likely larger.

Finally, our findings highlight two general lessons beyond OTC markets. First, markets that seem efficient might not be. With trade-level data, we find that the Canadian government bond market—which is liquid and offers a homogeneous, cash-like good—achieves only 60% of the first best. This is much lower than aggregate sufficient statistics that approximate the degree of efficiency (such as bid-ask spreads) would suggest. Second, introducing platforms to reduce frictions in decentralized markets—which are pervasive throughout the economy—has limitations. Platforms by Uber, Airbnb, Amazon, and others require sharing information, which consumers (here, investors) might dislike. Further, platforms might not be designed to foster competition, because they are owned by profit-maximizing firms (here, dealers).

(2015), p. 2).

Related literature. Our main contribution is to empirically assess price and welfare effects when OTC markets are centralized. Therefore, we add to relatively few studies that touch on the centralization of financial markets via reduced-form analysis (e.g., Barklay et al. (2006); Loon and Zhong (2016); Fleming et al. (2017); Abudy and Wohl (2018); Biais and Green (2019); Benos et al. (2020); Riggs et al. (2020); O’Hara and Zhou (2021)). Among them, the most closely related papers are Plante (2017) and Hendershott and Madhavan (2015) (HM). Plante (2017) studies whether to shift the bilateral trading of corporate bonds to a fully centralized, anonymous exchange. We focus on a less drastic market intervention.⁵ Similar to HM, we build a model in which investors choose between bilateral trading and trading on a multi-dealer platform; however, we highlight the trade-off dealers face when choosing quotes, which are exogenous in HM. Further, we structurally estimate our model, which allows us to conduct counterfactual analyses.

By creating a data set with trade-level information, we contribute to a steadily growing literature that analyzes trade-level data on financial markets (see Bessembinder et al. (2020) for an overview). Our market differs from those previously studied because it is highly liquid and features relatively high price transparency with little uncertainty about the true value of the asset.

By showing that even in this market there is evidence of price discrimination, we add to the descriptive evidence of price dispersion in less liquid or more opaque OTC markets (e.g., Green et al. (2007); Friewald and Nagler (2019); Hau et al. (2019)). Our findings differ from De Roure et al. (2020), who document an OTC discount in the market for German government bonds.

By estimating the demand and demand elasticity of an individual investor for government bonds, we contribute to a large literature that studies government bond markets using aggregate data (e.g., Garbade and Silber (1976); Krishnamurthy and Vissing-Jørgensen (2012, 2015)) and a young literature that estimates demand for

⁵Shifting bilateral trading onto an exchange (limit order book) is likely infeasible in the medium run, because regulators are reluctant to severely disturb the market. Further, empirical evidence suggests that investors might prefer auctions with dealers over limit order books (e.g., Riggs et al. (2020)).

financial assets (e.g., Koijen and Yogo (2019, 2020)). For estimation, we exploit techniques used to study (multi-unit) auctions to construct a cost-shifter instrument for prices outside of the auction (e.g., Hortaçsu and McAdams (2010); Kastl (2011); Hortaçsu and Kastl (2012); Allen et al. (2020)). Further, we apply an approach by Bresnahan (1981) that is commonly used in the literature on demand estimation to infer the marginal costs of firms from observable behavior in a trade setting. Here, marginal costs become values for realizing trade.

Our theory lies in between the theoretical literature on OTC markets (following Duffie et al. (2005)) and a large theoretical literature that studies decentralized or fragmented financial markets (with recent work by Glode and Opp (2019); Chen and Duffie (2021); Rostek and Yoon (2020); Wittwer (2021, 2020)). Similar to a few other papers, our model focuses on the selection of investors into trading venues (e.g., Liu et al. (2018); Vogel (2019)).⁶ Different from these papers, we highlight the importance of benchmark prices, as in Duffie et al. (2017), but we endogenize them.

Unlike most papers in the OTC literature, we do not highlight search frictions or price opaqueness, because the market we study is more liquid and price-transparent than other markets. This is similar to Babus and Parlato (2019), who study market fragmentation in OTC markets when there is no centralized platform, and to Baldauf and Mollner (2020), who show that it can be theoretically optimal for an investor to disclose information when running an auction. This is in line with our empirical findings.

Paper overview. The paper is structured as follows: Sections 2 and 3 describe the institutional environment and the data, respectively. Section 4 provides descriptive evidence that motivates the need for market reforms. Section 5 introduces the model, which is estimated in Section 6. Section 7 presents the estimation results that of the basis for the welfare analysis in Section 8, and Section 9 concludes.

⁶Following the literature, we assume exclusive participation per market segment. Dugast et al. (2019) only recently relaxed this assumption.

2 Institutional environment

Government bond markets are ideal for studying whether centralizing OTC markets can decrease dealer market power, leading in turn to welfare gains. The reason is that government bonds offer greater safety and liquidity than other securities. They are closer to a perfectly homogeneous good with a public market price and quick settlement. Therefore, we can rule out confounding factors that might drive markups and explain a decentralized market structure in other settings (such as high illiquidity, asymmetric information, counterparty risk, and product differentiation).

Market players. Government bond markets are populated by a few (in Canada, 10) primary dealers, which we refer to as dealers, and many investors. There are also smaller dealers and brokers, but they play a minor role in our case. Dealers are large banks, such as RBC Dominion Securities. Investors come in two types: they are either institutional or retail. Whether an investor is classified as institutional or retail is set by the Industry Regulatory Organization of Canada (see IIROC Rule Book⁷). The biggest classifying factor is how much capital an investor holds. Only if she holds enough does she qualify as an institutional investor.

To get a sense of who investors are, we manually categorize 1,459 investors we can identify by name. The largest investor groups are asset managers, followed by pension funds, banks, and firms that are members of IIROC. Many also work as asset managers. Then we have public entities (such as governments, central banks, and universities), insurance companies, firms that offer brokerage services, and non-financial companies (see Appendix Figure A1a).

Market structure. A country's government bond market often makes up a large part of its total bond market (in Canada 70%). It splits into two parts. The first is the primary market, in which the government sells bonds via auctions, primarily to dealers. The second and larger part is the secondary OTC market. It is similar to other OTC markets, with one segment in which dealers trade with other dealers (or brokers) and one in which dealers trade with investors. We focus on the larger (for

⁷Accessible at: www.iiroc.ca/industry/rulebook/Pages/default.aspx.

Canada) dealer-to-investor segment.

Trade realizes either via bilateral negotiation or on (an) electronic platform(s). These platforms are called alternative trading systems in the U.S. and Canada, multilateral trading facility in Europe, and dark pools for equities (see Bessembinder et al. (2020) for an overview). We focus on the most common type of platform in the dealer-to-investor segment, which matches investors to dealers but not to other investors.

Given the dealers' strong influence on OTC markets and the fact that it is not uncommon for dealers to own the platforms (as in Canada), there are reasons to believe that these platforms are not designed to maximize investor surplus. One indication of this is that dealers use a different kind of platform (anonymous limit order books) when trading with one another or with brokers from the one they use when trading with investors. Further, some platforms are only accessible to institutional investors.

In Canada, until recently, there was only one multi-dealer platform: CanDeal. It operates similarly to most other platforms (described below), including the largest ones in the U.S. and Europe.⁸ Yet unlike some platforms, CanDeal does not offer central clearing, different times to settlement, or higher price transparency than the bilateral market. This is useful for us, as it rules out confounding factors that might drive differences to bilateral trading.

How investors trade. If an investor is interested in trading a bond, she typically consults Bloomberg (or another information provider) to check dealer-advertised prices and prices that reflect the bond's current market value.

Next, the investor contacts her dealer, traditionally over the phone or by text. The dealer makes a take-it-or-leave-it offer that the investor either accepts or declines. If the investor declines, the investor could contact other dealers to seek more bilateral offers. A typical investor does not do so, however, as most investors have a single

⁸A non-exhaustive list of alternative trading systems includes MarketAxess (the leader in e-trading for global bonds), BGC Financial L.P. (which offers more than 200 financial products); BrokerTec Quote (which leads the European repo market); Tradeweb Institutional (global operator of electronic marketplaces for rates, credit, equities, and money markets).

dealer with whom they trade bilaterally (see Appendix Figure A1b).

As an alternative, institutional investors can trade on the platform CanDeal, where they have access to all dealers simultaneously. They can choose among different alternatives for how to trade, but the most common is to run a request for quote (RFQ) auction.

In an RFQ auction, an investor sends a request to up to four dealers. The request reveals to the dealers the name of the investor, whether it is a buy or sell, the security, the quantity, and the settlement date. Knowing how many—but not which—dealers are participating, dealers respond with a price. The investor chooses the deal that she likes best (typically the best price), and the trade is executed shortly after.

Running an RFQ auction differs from contacting multiple dealers bilaterally, because it is easier to make dealers compete. This is for three reasons. First, dealers see how many other dealers compete for the investor. Second, dealers have to respond simultaneously and not sequentially. This implies that the investor does not have to go back to a dealer she contacted earlier on, so the dealer cannot revise the offer and provide a worse quote. Third, running an auction is faster and requires less effort than contacting multiple dealers.

Platform usage costs. To use the platform, an eligible investor has to pay a small monthly fee, which ranges between C\$ 725 and C\$ 3,035, depending on usage. However, industry experts have raised concerns that the actual costs are indirect, because—despite platforms appearing to be an attractive alternative to bilateral trading—a relatively small fraction of trades actually occurs on platforms in many markets (e.g., McPartland (2016)). In our case, only about 35% of the institutional investors realize trades on the platform on a typical day.

These indirect costs might come from several sources. Switching to the platform rather than immediately realizing the trade with the dealer takes some time and effort. Investors could also be averse to sharing information about the trade with more than one dealer, or to damaging an existing relationship by appearing on the platform (Hendershott and Madhavan (2015); Managed Funds Association (2015);

Hendershott et al. (2020a)).⁹

3 Data

Our main data source contains trade-level information on all government bond trades of registered brokers or dealers. We augment this data with additional data sources.

Main data source. The main source is the Debt Securities Transaction Reporting System, MTRS2.0, collected by IIROC since November 2015. Our sample contains trade-level information on all bond trades of registered brokers or dealers from 2016 to 2019. The sample spans all trading days and 278 securities. We observe security identifiers (ISINs), the time, the side (buy/sell), the price, and the quantity of the trade. We also know whether an investor trades bilaterally or on the platform, and can identify whether the investor is institutional or retail as part of the reporting.

A unique feature of the data is that each dealer (and broker) carries a unique legal identifier (LEI). Investors either have an LEI or a dealer-specific identifier. The latter is an anonymous dealer-specific account ID. Of all trades with investors, about 25% are with investors that have an LEI. These can be identified by name and traced throughout the market.

Similar to the TRACE data set, the MTRS2.0 data are self-reported and requires cleaning (see Appendix A). The cleaned sample includes almost all (cash) trades of Canadian government bonds but misses trades between investors, which do not have to be reported but are rare, according to market experts.¹⁰ To get a sense of how

⁹Relationships might matter for different reasons. For instance, the investor may be able to post low collateral with the dealer to trade (low margins), may have better options of realizing a potentially large trade fast and at a good price when she needs it (balance sheet space), or get better terms for other activities, such as lending or borrowing overnight (haircuts). Finally, an investor might simply like to be told that she obtains the most favorable price, thanks to her loyalty.

¹⁰In line with this, participation on all-to-all platforms—on which investors can directly trade with one another—remains low in markets in which these platforms already exist (Bessembinder et al. (2020)). One example is the U.S. government bond market, as was discussed at the 2020 U.S. Treasury Market Conference.

many trades our sample misses because of this or due to misreporting, we compare the daily trading volume of Treasury Bills in MTRS2.0 with the full volume, which must be reported to the Canadian Depository for Securities. Our data cover approximately 90% of all trades involving Treasury Bills.

Additional data sources. We augment our trade-level data with additional data sources. First, we obtain bidding data on all government bond auctions between 2016 and July 2019 from the Bank of Canada. We can see who bids (identified by LEI) and all winning and losing bids. Importantly, we can link how much a dealer won in the primary market to how she trades in the OTC market, which we use to construct an instrument in our demand estimation.

Second, we scrape ownership information from the public registrar of LEIs (gleif.org). This tells us whether a counterparty LEI is a subsidiary of a dealer so that we can exclude in-house trades, which are between a dealer and one of its subsidiaries.

Lastly, we collect averages of the (indicative) quotes posted on Bloomberg (BNG) and on the platform (CanDeal) for each security per hour. Of these quotes, the Bloomberg mid-quote (which is the average between the bid and the ask quote) is the most important. We use this price as a proxy for a bond's true market value, which is commonly known by everyone. We believe that this is a reasonable assumption, because the BNG mid-price is very close to the price at which dealers trade with one another. The inter-dealer market price, in turn, is often taken as the true value of a security in the related literature.

Sample restrictions. We exclude in-house trades because they are likely driven by factors that differ from those of a regular trade; for instance, tax motives or distributing assets within an institution. In addition, we exclude trades that are realized outside of regular business hours (before 7:00 am and after 5:00 pm), because these trades are either realized by foreign investors who might be treated differently or by investors who are exceptionally urgent to trade.

For the estimation of our structural model, we impose some additional restriction in order to construct an instrument for quotes using bidding data on the primary auctions. We focus on primary dealers and drop trades after July 2019 because we

do not have auction data for the second half of 2019. Due to data reporting, we exclude one dealer. Further, we exclude trades that are realized before the outcome of a primary auction was announced—10:30 am for bill auctions and 12:00 am for bond auctions. Appendix Table 1 summarizes all sample restrictions.

3.1 Key variables and market features

Unit of measurement. Following market conventions, we convert each price into the yield-to-maturity (the annualized interest rate that equates the price with the present discount value of the bond) and report our findings in terms of yields rather than prices; a higher price implies a lower yield, and vice versa.

All yields are expressed in bps; 1 bps is 0.01%. This is a relatively large yield difference because of the low interest rate level throughout our sample. As comparison, the median yield of a bond is about 150 bps, and the median bid-ask spread is about 0.5 bps.

Normalization. The yield (and price) of a bond might be affected by many factors, and explaining all of them in a single model is beyond the scope of this paper. Our approach, instead, is to control for factors that are not endogenous in our model (as discussed in Section 5.3). We do this by regressing the $yield_{thsi j}$ of a trade on day t in hour h of security s between dealer j and investor i on an indicator variable that separates trades in which the investor buys from trades in which she sells, a flexible function of trade size, $f(quantity_{thsi j}) = \sum_p^3 (quantity_{thsi j})^p$, an hour-day fixed effect, and a security-week fixed effect.¹¹ We then construct the residual from this regression. In addition, we normalize the Bloomberg yield by subtracting the estimated hour-day and security-week fixed effects.

We label the residualized trade yield $y_{thsi j}$ and the normalized Bloomberg yield θ_{ths} . For consistency, we use these normalized yields throughout the paper—that is, for the reduced-form evidence as well as to estimate our model. However, our reduced-form evidence is robust to using the yields in the raw data rather than the normalized yields.

¹¹Our results are not sensitive to how we control for trade size.

Key market features. The typical trade between a dealer and one of the 546,048 investor IDs is small (see Appendix Figure A3). It involves a bond that is actively traded, because it was issued in a primary auction less than three months ago and is on-the-run. Bid-ask spreads are narrow (0.5 bps at the median), and it takes only 0.13 (2.8) minutes between an investor who buys and an investor who sells (the same security) on a day. Taken together, the market is highly liquid.

4 Descriptive evidence: Why market reforms?

Our trade-level data allow us to document yield differences across investors, which suggests that the market is imperfect.

Markups and yield gap. To analyze whether dealers charge markups over the market value (θ_{ths}), we define the markup as:

$$(y_{thsi j} - \theta_{ths})^+ = \begin{cases} y_{thsi j} - \theta_{ths} & \text{when the investor buys} \\ \theta_{ths} - y_{thsi j} & \text{when the investor sells.} \end{cases} \quad (1)$$

The higher $(y_{thsi j} - \theta_{ths})^+$, the more favorable the yield for the investor, independent of whether she buys or sells.

Figure 1 shows that markups vary widely across investors, even when controlling for differences in trade size, security ID, time of trade, and dealer. Furthermore, these markups are systematically smaller for institutional investors—who have access to the platform—than for retail investors, who do not. At the median, a retail investor obtains a yield that is about 4 bps worse than an institutional investor. This amounts to a yearly monetary loss of roughly C\$ 34,000 for the average retail investor who trades C\$ 86 million per year.

Whether market reforms that shift trading onto the platform could have sizable effects on yields or welfare depends on what drives the yield gap. It could be that retail investors are willing to pay more and therefore realize worse yields. But it could also be that they obtain worse yields because they cannot trade on the platform. In that case, there is scope for market reforms.

Figure 1: Yield gap

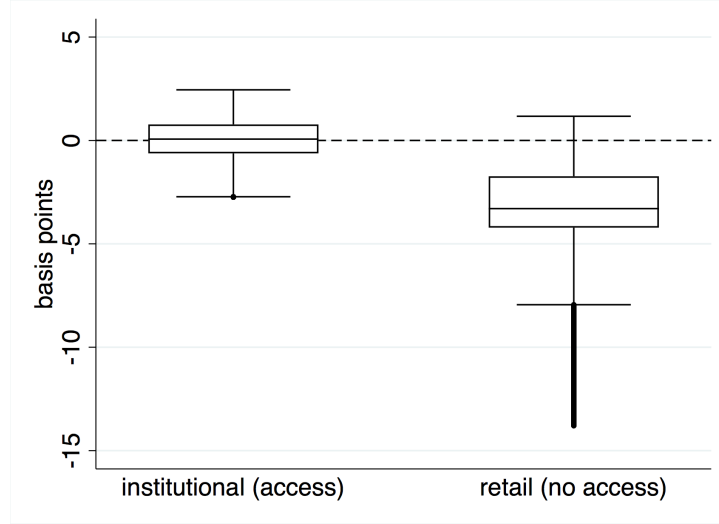


Figure 1 shows a box plot of markups for institutional and retail investors, excluding the upper and lower 5% of the distribution. To construct these markups, we regress $(y_{thsi j} - \theta_{ths})^+$ as defined in (1) on indicator variables that distinguish retail from institutional investors ($retail_{thsi j}$) and buy-side from sell-side trades ($buy_{thsi j}$). In addition, we control for hour-day (ζ_{th}), security (ζ_s), and dealer (ζ_j) fixed effects. The residual measures how much worse the yield is relative to the market value.

Yields and platform access. The ideal but infeasible experiment to establish a causal link between platform access and yields would be to randomly assign platform access to some investors (treatment group) but not all investors (control group). Instead, we leverage the fact that 90 institutions lost the right to access the platform in our sample to conduct an event study.

Investors lose platform access when they lose their institutional status. This can happen for different reasons. First, the investor may no longer hold sufficient capital or may no longer be willing to prove that she does. For instance, the non-financial or financial assets of a firm could lose value so that a firm no longer has a net worth of C\$ 75 million, which is the cutoff to classify as institutional investor. Second, the investor may terminate her membership with a regulated entity such as the Canadian Investor Protection Fund (CIPF), which protects investor assets in case of bankruptcy. Third, the investor may stop selling securities, offering investment advice, or managing a

mutual fund. Our data do not allow us to disentangle these reasons because switchers are not reported with their LEI, and are therefore anonymous to us.

We define the event of investor i as the first time we observe this investor as a retail investor. We bucket time into months and pool the buy and sell side of the trade to obtain sufficient statistical power. We then test whether the investor obtains a worse yield ($y_{thsi j}$) relative to the market value (θ_{ths}) when losing platform access by regressing

$$(y_{thsi j} - \theta_{ths})^+ = \zeta_i + \sum_{m=Mi^-}^{Mi^+} \beta_m D_{mi} + \zeta_{th} + \zeta_s + \zeta_j + \epsilon_{thsi j}, \quad (2)$$

where D_{mi} is an indicator variable equal to 1 m months before/after i loses access and $\zeta_i, \zeta_{th}, \zeta_s, \zeta_j$ are investor, hour-day, security, and dealer fixed effects, respectively.

Our parameters of interest (the β 's) are identified from how the trade yields of an investor who loses access change over time when controlling for time, security, and dealer-specific unobservables. The sizes of the hour, security, and dealer fixed effects are pinned down by trade information on retail investors who never obtain access (in our sample), since these investors are likely more similar to those who lost access than to institutional investors. This is because investors with access throughout tend to be large players and clearly meet the regulatory requirements.

We find that investors who lose platform access realize worse yields (see Figure 2). The yield drops on average by 1 bps in the first month and decreases further by about 4 bps thereafter. This suggests that platform access matters for yields.

The relationship would be causal if losing access to the platform were exogenous. This would be the case if an investor's regulatory capital fell below the regulatory threshold due to shocks unrelated to trading demand. In this case we would expect a change in investor-status but not a change in trading behavior. Alternatively, an investor might lose the status for potentially endogenous reasons, such as a change in business model. Then we would expect to see changes in trading behavior. In Appendix Figure A2 we show that investors who lose access do not systematically change trading behavior after they lose access—providing us confidence that losing access is exogenous.

Figure 2: Event study: Yield drop when losing platform access

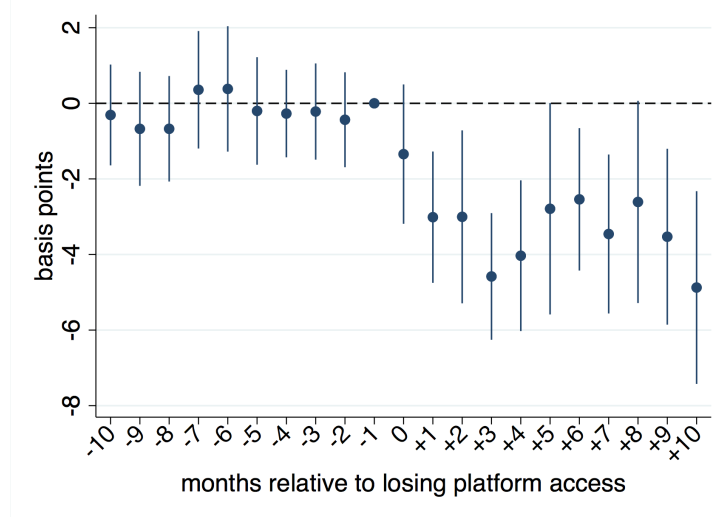


Figure 2 shows the β_m estimates and the 95% confidence intervals of event study regression (2) for 10 months before and after an investor i switched from having to not having access. Each β_m measures by how much the markup, $(y_{thsi,j} - \theta_{ths})^+$, for investor i differs m months before/after this event relative to 1 month prior to it. Standard errors are clustered at the investor level.

Summary. We have gathered novel evidence that platform access matters for yields. This implies that there is scope to increase investor yields by centralizing the market.

5 Model

We now introduce a model to gain additional insights. The model allows us to quantify total gains from trade, our measure of welfare. For this we need to know the values for realizing trade. In the data, we only observe transaction prices, which lie somewhere between the value of the buyer and the value of the seller. Further, we can account for institutional investors selecting between trading bilaterally and on the platform. This selection problem can bias estimates of OLS regressions. Lastly, we can take into account how dealers and investors respond to changes in market structure, such as centralizing trades on a platform.

Without loss of generality, we consider two separate games: one in which dealers

sell to investors and one in which dealers buy from investors.¹² We explain the setting with buying investors; the other side is analogous. We use yields rather than prices in line with the rest of the paper and estimation. To make the price-yield conversion, it helps to keep in mind that the yield is like a negative price. We denote a vector of quotes by $q_t = (q_{t1} \dots q_{tJ})$ and similarly for all other variables. Random variables are highlighted in **bold**. All proofs are in Appendix B.1. Simplifying assumptions are discussed in Section 5.3.

5.1 Model overview

Dealers sell a bond to institutional and retail investors in a two-stage game, which is inspired by how investors trade in this market (see Section 2).

First, dealers simultaneously set quotes to maximize the expected profit from trading with investors. The quotes are posted publicly and inform investors about the yields they may expect to realize when buying on the platform. This is motivated by the empirical fact whereby individual trade yields on the platform are on average identical to the quotes dealers post on the platform. Then, given these quotes, institutional investors decide whether to buy on the platform or bilaterally with a dealer, while retail investors can only buy bilaterally.¹³

In choosing a quote, a dealer faces a trade-off: On the one hand, when the dealer decreases the quote, the platform becomes less attractive for investors. As a result, more investors stay off the platform and buy bilaterally. On the other hand, when she increases the quote, more investors entering the platform buy from her, increasing her platform market share.

¹²To see why considering separate games is without loss of generality, assume that investors may either buy or sell, but that the dealer does not know whether the investor is a buyer or seller. The dealer offers a bid and ask yield, such that the bid is optimal conditional on the investor's being a seller and the ask is optimal conditional on the investor's being a buyer. The ask (bid) yield is identical to the ask (bid) yield in our model with only buying (selling) investors.

¹³We assume that all institutional investors consider the platform to be an alternative because we observe no systematic difference between investor types who trade on versus off the platform (see Appendix Figure A1a).

An institutional investor also has a trade-off: When she buys bilaterally she has to leave all surplus to the dealer because the dealer discovers her willingness to pay. When entering the platform the investor can extract positive surplus thanks to (more direct) competition between dealers but has to pay a cost to use the platform. As a result, in equilibrium only investors with a high willingness to pay enter the platform.

5.2 Formal details

On a fixed day t , $J_t \geq 2$ dealers sell a bond to infinitely many investors, bilaterally or on a platform. Each transaction is a single unit trade. The market value of the bond is $\theta_t \in \mathbb{R}^+$. It is commonly known and exogenous.

Dealers and investors. Motivated by the empirical feature that investors tend to trade with a single dealer bilaterally, each investor i has a home dealer d , short for d_i . Each dealer, thus, has a home investor base. It consists of two investor groups, institutional and retail investors, indexed by $G \in \{I, R\}$. Each has a commonly known mass κ^G of potential investors. W.l.o.g., we normalize $\kappa^I + \kappa^R = 1$. Of the potential investors in group G , N_t^G investors actually seek to buy on any particular day. This number is exogenous and unknown to the dealer until the end of the day.

Each dealer seeks to maximize profit from trading with investors. Ex post, dealer j obtains a profit of $\pi_{tj}(y) = v_{tj}^D - y$ when selling one unit at yield y . Here, $v_{tj}^D \in \mathbb{R}$ is the dealer's value for the bond. It may be driven by current market conditions, expectations about future demand, or prices or inventory costs.

If the market was frictionless and dealers neither derived value from holding bonds nor paid any costs for intermediating trades, v_{tj}^D would equal the market value, θ_t . However, since this is unlikely in reality—for instance, because it is costly to hold inventory—we refrain from imposing $v_{tj}^D = \theta_t$.

An investor $i \in G$ obtains a surplus of $y - v_{tid}^G$ when buying at yield y , where

$$v_{tid}^G = \theta_t + \nu_{ti}^G - \xi_{td} \text{ with } \nu_{ti}^G \stackrel{iid}{\sim} \mathcal{F}_t^G \text{ and } \xi_{td} \sim \mathcal{G}_t$$

is the investor's value, also referred to as willingness to pay. It splits into three elements. The first is the commonly known market value of the bond, θ_t . The second is a privately known liquidity shock, ν_{ti}^G , which is drawn iid from a commonly known

distribution with a continuous CDF $\mathcal{F}_t^G(\cdot)$ that has a strictly positive density on the support. It reflects individual hedging or trading strategies, balance-sheet concerns, or the cash needs of an institution. The third element is dealer-specific, ξ_{tj} . It is drawn from an arbitrary distribution with CDF $\mathcal{G}_t(\cdot)$ and absorbs unobservable dealer characteristics (similar to a fixed effect in a linear regression). We label this term dealer quality, as it captures anything that makes trading with a specific dealer particularly attractive, independent of how the trade is realized. It could, for example, reflect the dealer’s probability of delivery, the speed of processing the trade, or her ability to hold or release large quantities or to provide ancillary services (such as offering investment advice on a broad range of securities).

Timing of events. The game has two stages. In the first stage, dealers simultaneously post indicative quotes at which they are willing to sell on the platform, which is accessible to institutional investors. When choosing the quote, each dealer maximizes the expected profit from selling to institutional investors. Each supposes that they can sell on the platform at the posted quote and forms expectations over how much institutional investors are willing to pay.

In the second stage of the game, all trades realize. In a bilateral trade, the dealer discovers the investor’s willingness to pay and offers a yield equal to that.¹⁴ On the platform, each investor runs an auction with the dealers; this determines by how much their individual platform trade yield differs from the posted quotes.

We formalize such an auction-game in Appendix B.3.¹⁵ Here we only give the main idea, because we cannot estimate the auction-game without bidding data from the platform. Before bidding in the auction for investor i , each dealer j draws a signal,

¹⁴This implies that the dealer may occasionally accept to trade at a loss. This happens when the investor draws an extreme liquidity shock that lies above the dealer’s value, and captures the idea that a dealer is willing to occasionally help an investor in need to sustain their bilateral relationship.

¹⁵Our auction-game is a possible micro-foundation for why trade yields relate to quotes, as in equation (3). As an alternative, we could formalize the fact that a dealer might not respond in an RFQ auction (as in Liu et al. (2018); Riggs et al. (2020)).

ϵ_{tij} , that updates her value.¹⁶ It is drawn from a commonly known distribution with CDF $\mathcal{H}_t(\cdot)$. Each dealer then offers a bid that may differ from her posted quote, q_{tj} , for a given realization of the signal. Yet to avoid having the quote being seen as cheap talk, the bid must be proportional to the quote in expectation. We show that in equilibrium, each dealer's bid is equal to her posted quote plus a stochastic term that depends on the signal, ϵ_{tij} , and a parameter σ :

$$q_{tj} + \sigma \epsilon_{tij} \text{ where } \epsilon_{tij} \sim \mathcal{H}_t \text{ and } \sigma \in \mathbb{R}. \quad (3)$$

Parameter σ measures the degree of competition on the platform and depends, for instance, on the number of dealers who bid in the auction.¹⁷ When $\sigma = 0$, all investors buy from the dealer who offers the best quote and quality, i.e., the one with $\max_j \{\xi_{tj} + q_{tj}\}$. In that case, the platform is perfectly competitive and dealers compete à la Bertrand. If $\sigma \rightarrow \infty$, investors buy from the dealer for which the realization of ϵ_{tij} is the highest, regardless of the dealer's quote or quality. In this case, each dealer acts as a monopolist on the platform.

To use the platform and choose from any of the dealers, the investor has to pay a commonly known cost c_t . This represents any obstacle to access the platform, including privacy concerns or relationship costs, and is motivated by the empirical fact that even though platform yields are better than bilateral yields (see Appendix Table 2), on a typical day only 35% of institutional investors use the platform. The cost also absorbs differences in the service a dealer provides on versus off the platform. Although uncommon, a bilateral trade could be part of a package or come with additional investment advice.

In summary, the sequence of events is:

- (0) Dealers observe $\theta_t, \xi_t, v_{tj}^D$, and investors observe θ_t .
 N_t^G investors of both groups $G \in \{I, R\}$ draw liquidity shocks ν_{ti}^G .
- (1) Dealers simultaneously post $q_{tj} \in \mathbb{R}^+$ for everyone to see.

¹⁶More broadly, in our estimation, ϵ_{tij} can capture anything that prevents an investor from buying from the best dealer with the highest posted quote.

¹⁷More broadly, in our estimation, σ could be shaped by a variety of factors, including limited information- or risk-sharing (as in Boyarchenko et al. (2021)).

- (2) Each investor contacts her home dealer, who observes ν_{ti}^G and offers $y_{tid}^G = v_{tid}^G$. Retail investors accept the offer. An institutional investor can accept or enter the platform. In the latter case, the investor pays cost c_t , observes the platform shock ϵ_{tij} , and decides from which dealer to buy at $q_{tj} + \sigma\epsilon_{tij}$.

A pure-strategy equilibrium can be derived by backward induction.

Proposition 1 (Investors).

- (i) A retail investor with shock ν_{ti}^R buys bilaterally from home dealer d at

$$y_{tid}^R = \theta_t + \nu_{ti}^R - \xi_{td}. \quad (4)$$

- (ii) An institutional investor with shock ν_{ti}^I buys bilaterally from home dealer d at

$$y_{tid}^I = \theta_t + \nu_{ti}^I - \xi_{td} \quad \text{if } \psi_t(q_t) \leq \nu_{ti}^I, \quad (5)$$

$$\text{where } \psi_t(q_t) = \mathbb{E}[\max_{k \in \mathcal{J}_t} \tilde{u}_{tik}(\boldsymbol{\epsilon}_{tik})] - \theta_t - c_t \text{ with } \tilde{u}_{tij}(\boldsymbol{\epsilon}_{tij}) = \xi_{tj} + q_{tj} + \sigma\boldsymbol{\epsilon}_{tij}. \quad (6)$$

Otherwise, the investor enters the platform, where she observes ϵ_{tij} and buys from the dealer with the maximal $\tilde{u}_{tij}(\epsilon_{tij})$ at $q_{tj} + \sigma\epsilon_{tij}$.

This proposition characterizes where investors buy and at what yields. A retail investor always buys at a yield that equals her willingness to pay (statement (i)). An institutional investor trades on the platform if she expects the surplus from buying on the platform minus the platform usage cost, $\mathbb{E}[\max_{k \in \mathcal{J}_t} \tilde{u}_{tik}(\boldsymbol{\epsilon}_{tik}) - (\theta_t + \nu_{ti}^I)] - c_t$, will be higher than the zero surplus she receives in a bilateral trade (statement (ii)). This is the case for urgent investors who are willing to pay a higher price, i.e., accept a low yield due to a low liquidity shock. For them it is better to trade on the platform, because the platform quote is targeted to an investor with an average willingness to pay rather than to the investor's individual willingness to pay.

The proposition highlights the fact that yields for institutional investors are higher than for retail investors because they have access to the platform: Those who obtain better yields on the platform buy on the platform; others buy bilaterally.

Proposition 2 (Dealers). Dealer j posts a quote q_{tj} that satisfies

$$q_{tj} \left(1 + \frac{1}{\eta_{tj}^E(q_t)} \left(1 - \frac{\partial \pi_{tj}^D(q_t)}{\partial q_{tj}} / S_{tj}(q_t) \right) \right) = v_{tj}^D, \quad (7)$$

where $\eta_{tj}^E(q_t)$ is the dealer's yield elasticity of demand on the platform and $\frac{\partial \pi_{tj}^D(q_t)}{\partial q_{tj}}$ is the marginal profit the dealer expects from bilateral trades with institutional investors. It is normalized by the size of the dealer's platform market share, $S_{tj}(q_t)$. Formally, $\eta_{tj}^E(q_t) = q_{tj} \frac{\partial S_{tj}(q_t)}{\partial q_{tj}} / S_{tj}(q_t)$ with $S_{tj}(q_t) = \sum_{k \in J_t} \Pr(\boldsymbol{\nu}_{\mathbf{ti}}^I \leq \psi_t(q_t)) \Pr(\tilde{u}_{tki}(\boldsymbol{\epsilon}_{\mathbf{tki}}) < \tilde{u}_{tij}(\boldsymbol{\epsilon}_{\mathbf{tij}}) \forall k \neq j)$, where $\tilde{u}_{tij}(\boldsymbol{\epsilon}_{\mathbf{tij}})$, $\psi_t(q_t)$ as in (6), and $\pi_{tj}^D(q_t) = \mathbb{E}[v_{tj}^D - (\boldsymbol{\nu}_{\mathbf{ti}}^I + \theta_t - \xi_{tj}) | \psi_t(q_t) \leq \boldsymbol{\nu}_{\mathbf{ti}}^I]$.

Proposition 2 characterizes the quotes dealers post on the platform. Taking the quotes of the other dealers as given, each dealer chooses a quote that equals a fraction of her value, v_{tj}^D . To obtain an intuition regarding what determines the size of this fraction, it helps to abstract from the bilateral segment for a moment.

If the market consisted of the platform only, the dealer's quote would satisfy $q_{tj}(1 + 1/\eta_{tj}^E(q_t)) = v_{tj}^D$. This is equivalent to the classic markup rule of firms that set prices to maximize profit. Each chooses a price that equals its marginal cost multiplied by a markup, which depends on the price elasticity of demand, $\eta_{tj}^E(q_t)$. In our setting, the marginal cost is v_{tj}^D , and since the dealer chooses quotes in yields rather than prices, the markup is actually a discount.

When the market splits into the platform and a bilateral segment, there is an additional term. It captures the fact that a quote also affects how much profit the dealer expects to earn from bilateral trades, given that investors select where to buy based on these quotes. If the dealer decreases the quote, more investors buy bilaterally because they earn a higher yield there; how many depends on the cross-market (segment) elasticity between bilateral and platform trading. If this elasticity is high, investors easily switch onto the platform. To prevent this from happening, the dealer decreases the quote to make the platform less attractive.

In summary, when choosing the quote the dealer trades off the profit from selling bilaterally, where she extracts a higher trade surplus, with the profit she earns on the platform when stealing investors from other dealers.

5.3 Discussion

Our model builds on several simplifying assumptions. First, because we do not observe failed trades, we assume that the number of dealers and investors who trade on a day is exogenous and that no trade between them fails. We believe that this is not problematic for two reasons. First, empirical evidence suggests that trades of safe assets rarely fail (e.g., Riggs et al. (2020); Hendershott et al. (2020b)). Second, (primary) dealers have an obligation to actively trade: The least active dealer trades on 98% of dates. We can, thus, abstract from market entry and exit of dealers.

Second, our game does not connect multiple days. In particular, we assume that dealers' and investors' values for the bond are independent of prior trades. This implies that we set aside dynamic trading strategies. Dealers and investors can still trade every day and their values can capture continuation values, which may vary in time. However, when changing the market rules, we cannot account for changes in their continuation values.

Third, we abstract from an inter-dealer market, because dealers primarily trade with investors in our data. In particular, it is not the case that dealers perfectly balance out their inventory positions by trading with one another. Instead, they carry significant amounts of bonds in inventory, which can explain why different dealers have different values for buying or selling.

Fourth, in our main specification we assume that the dealer offers a bilateral yield that equals the investor's full willingness to pay, leaving the investor with zero trade surplus.¹⁸ We do this because there is no information in the data that allows us to identify a dealer's beliefs about how much investors are willing to pay or their bargaining power. From a theoretical viewpoint, this is a relatively strong assumption. It implies that dealers do not adjust the yields they charge in bilateral trades depending on how costly it is for them to realize the trade. We test whether this

¹⁸An alternative would be to assume that the dealer sets the bilateral yield in order to render an institutional investor indifferent between trading bilaterally or entering the platform. If platform entry is costly and bilateral trading is the default, no one would enter the platform in equilibrium. If platform entry is free, bilateral yields should be weakly better than platform yields. This is not supported by the data.

implication holds in our data and find supporting evidence (see Appendix Table 5). In addition, we show that our findings are robust to allowing investors to extract a fraction $\phi \in [0, 1]$ of surplus in the bilateral trade. For this we extend the model to incorporate a Nash-bargain for the yield in Appendix B.2.

Fifth, given that most trades in our sample are small and of similar size, we assume that all trades have the same size, normalized to one (see Appendix Figure A3).¹⁹ This implies that our findings are expressed in terms of unit of the trade and that we cannot analyze whether and how trade sizes and volume change as we change the market rules.

Finally, we abstract from order splitting by assuming that investors trade either bilaterally or on the platform but not both, because we observe that investors typically maximally trade once per day and do not split orders (see Appendix Figure A4).

6 Estimation

Including both the buy and sell sides of the market, we have four investor groups, indexed by G : retail and institutional investors who buy (R and I) and sell (R^* and I^*), respectively. For all of them, we want to estimate the daily distribution of the liquidity shocks (F_t^G), in addition to the daily dealer qualities (ξ_{tj}) and the degree of competition on the platform (σ). We allow the cost of using the platform and the dealer's value to depend on whether the investor buys (c_t and v_{tj}^D) or sells (c_t^* and v_{tj}^{*D}).

Notice that most of the parameters are day-specific. This allows us to nonparametrically account for variation and correlation across days that are driven by unobservable market trends. This is important because the yields and demands for Canadian government bonds are largely affected by global macroeconomic trends. To obtain sufficient statistical power, we also cannot allow the platform usage cost or the taste for dealer quality to be heterogeneous across investors. Instead, we estimate both for the average investor. Details of the estimation are in Appendix C.

¹⁹Appendix Tables 3 and 4 provide additional supporting evidence. In the related theoretical literature, following Duffie et al. (2005), this assumption is common.

6.1 Identifying assumptions

Our estimation builds on four identifying assumptions, a normalization and two parametric assumptions that are not crucial for identification.

Assumption 1. *Within a day t , the liquidity shocks ν_{it}^G are iid across investors i in the same group G .*

This assumption would be violated if an investor trades more than once in a day and jointly decides whether and at what price to trade for all such trades. However, given that we observe very few investors who trade several times within the same day, this is unlikely (see Appendix Figure A4).

Assumption 2. *W.l.o.g. we decompose dealer quality, ξ_{tj} , into a part that is persistent over time and a part that might vary: $\xi_{tj} = \xi_j + \chi_{tj}$. The time-varying parts, χ_{tj} , capture day- and dealer-specific demand shocks that are unobservable to the econometrician. They are drawn iid across dealers j within a day t .*

Crucially, Assumption 2 does not rule out that dealers may change quotes in response to demand shocks that are unobservable to the econometrician. Formally, ξ_{tj} may be correlated with q_{tj} . To eliminate the implied endogeneity bias, we need an instrument for the quotes.

Our solution is to extract unexpected supply shocks, $won_{\tilde{t}j}$, from bidding data in the primary auctions in which the government sells bonds to dealers:

$$\begin{aligned} won_{\tilde{t}j} = & \text{amount dealer } j \text{ won at the last auction day } \tilde{t} \\ & - \text{amount she expected to win when placing her bids.} \end{aligned} \quad (8)$$

These shocks work as cost-shifter instruments. This is because it is cheaper for dealers to satisfy investor demand when unexpectedly winning a lot at auction, given that auction prices are systematically lower than prices in the OTC market.

Importantly, we use the expected rather than the actual winning amounts, since dealers anticipate or even know investor demand when bidding in the auction—for example, because investors place orders before and during the auction (as in Hortaçsu and Kastl (2012)). This information affects how dealers bid and, consequently, how

much they win, which creates a correlation between the unobservable demand shocks and the actual, but not expected, winning amounts.

To compute the expected amount and control for anything the dealer knows at the moment she places her bids, we model the bidding process in the auction and use techniques from the empirical literature on (multi-unit) auctions. In a nutshell, we fix a dealer in an auction, randomly draw bids (with replacement) from the other bidders, and let the market clear. This generates one realization of how much the dealer wins. Repeating this many times generates the empirical distribution of winning amounts, from which we compute the expectation (see Appendix C.1 for details).

Assumption 3. *Conditional on unobservables that drive aggregate demand and supply on day t , ζ_t , and the time-invariant quality of the dealer, ξ_j , the demand shocks, χ_{tj} , are independent of the unexpected supply shocks, won_{tj} : $\mathbb{E}[\chi_{tj} | won_{tj}, \zeta_t, \xi_j] = 0$.*

To better understand whether this assumption is plausible, it helps to think through where the surprise—and, with that, the identifying variation—comes from. For one, the dealer is surprised when the Bank of Canada issues a different amount to bidders than the dealer expected. However, the date fixed effect absorbs most of this effect. What is left is the surprise the dealer faces when other bidders bid differently than the dealer expected.

With this in mind, the biggest threat to identification is the following scenario: One dealer is hit by a negative shock and bids less, so that the other dealers win more than expected. If investors substitute from the unlucky dealer toward those who won more, the exclusion restriction would be violated. However, in our data, we see relatively little substitution of investors across dealers. Therefore, we are less worried that this is a first-order concern.

The exclusion restriction would also be violated if the dealer changed her quality based on how much she won at auction. We believe that this is unlikely to happen (often) for at least two reasons. First, dealers have incentives to smooth out irregular shocks to maintain their reputation and their business relationships with investors in the longer run. Second, dealers would risk revealing information about their current inventory positions if they changed the service they provide based on how much they

win at auction.

Assumption 4. *Platform shocks ϵ_{tij} are iid across t, j, i .*

This assumption is frequent in demand estimation. It implies that the model restricts how investors substitute across different dealers. Since this type of substitution is rare in the data, and hence not the focus of this paper, we do not believe this assumption is problematic.

We normalize the quality of one dealer to 0, because—as is common in demand estimation—we cannot identify the size of the dealers’ qualities but we can estimate the quality differences between dealers.

Normalization 1. *The benchmark dealer ($j = 0$) provides zero quality: $\xi_{t0} = 0 \forall t$.*

Finally, we rely on two parametric assumptions. The first imposes a functional form on the distribution of ϵ_{tij} that is standard in demand estimation. It implies that the dealer’s market shares (on the platform) have a closed-form solution. The second assumption is inspired by the shape of the histogram of shocks, $\hat{\nu}_{ti}^G = y_{tij}^G - \theta_t + \hat{\xi}_{tj}$, for investors who choose to trade bilaterally. It resembles a normal distribution, similar to Figure 4.²⁰

Parametric Assumptions.

- (i) *Platform frictions ϵ_{tij} are extreme value type 1 (EV1) distributed.*
- (ii) *Liquidity shocks ν_{ti}^G are drawn from a normal distribution $N(\mu_t^G, \sigma_t^G)$ for all g, t .*

6.2 Identifying variation

We estimate the model separately for each investor group. Here, we focus on buying institutional investors and leave the other groups for Appendix C.2.

²⁰Without imposing a distributional assumption on the liquidity shocks, we could nonparametrically estimate the (truncated) distribution of the liquidity shocks of investors who trade bilaterally—as explained in more detail below—as well as bounds on the cost parameters.

Key variables. The first set of variables used in the estimation are each dealer’s daily market share on the platform (among investors who enter the platform), s_{tj} , and each dealer’s daily bilateral market share relative to her platform market share, ρ_{tj} . The second set of variables are the normalized yields of Section 3.1. We approximate the bond’s daily market value, θ_t , by the normalized Bloomberg yield, averaging across securities and hours of the day. Further, we approximate the quote at which dealer j sells on a day, q_{tj} , by the average yield at which she sells on the platform on that day. We believe this is a reasonable approximation because the posted average quote (across dealers) we observe is very similar to the average of the trade yields on the platform.

Identification. The main identifying variation for the competition parameter and the dealers’ qualities comes from how dealers split the platform market on a day.

The competition parameter (σ) is mainly identified from the within-day correlation between dealers’ daily platform market shares and their (cost-shifter) supply shocks (see Figure 3). To derive an intuition for this, assume for a moment that dealers do not differ in quality ($\xi_{tj} = 0 \forall j$). If the platform is perfectly competitive ($\sigma = 0$), a single dealer—namely the one with the most favorable supply shock and with it the best quote q_{tj} —captures the entire platform market share on that day. As σ increases, this dealer loses more and more of her market share to the other dealers. How much of the market share each dealer gains depends (besides σ) on the dealers’ supply shocks. Hence, the correlation between these market shares and the supply shocks pins down σ .

The dealers’ qualities (ξ_{tj}) are determined by how the dealers split the platform market when posting the same or very similar quotes: Dealers with higher qualities capture a higher market share.²¹

The distribution of the liquidity shocks and the platform usage costs are, for any given day, mainly identified from how bilateral yields vary across investors and

²¹We expect dealer quality to vary across dealers because we observe systematic differences in how much of the market each dealer captures when posting the best quote (see Figure 6a).

Figure 3: Dealers' daily platform market shares and their cost shifters

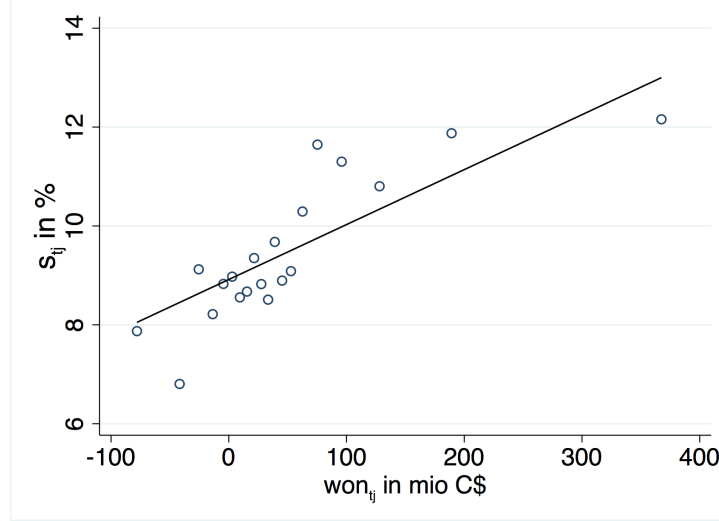


Figure 3 shows a binned scatter plot to visualize the correlation between dealers' daily market shares on the platform (s_{tj}) and their unexpected supply shocks ($won_{\tilde{t}j}$) when partialling out day fixed effects.

Figure 4: Identifying variation for c_t, μ_t^I, σ_t^I

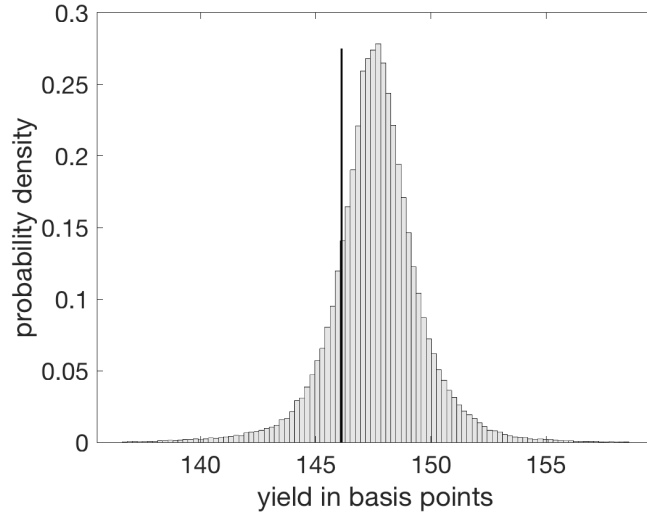


Figure 4 shows a probability density histogram of the yields (in bps) that institutional buyers realize, excluding the upper and lower 0.1 percentile of the distribution, and a black line. This line is the average cutoff (across days and dealers) that determines whether an institutional investor buys bilaterally or on the platform, according to Proposition 1.

Table 1: Estimates (median across days)

Buys	$\hat{\mu}^I$	$\hat{\mu}^R$	$\hat{\sigma}^I$	$\hat{\sigma}^R$	\hat{c}	$1/\hat{\sigma}$	$\hat{\eta}$
	-0.82 (0.13)	-2.92 (0.75)	2.81 (0.10)	5.12 (0.94)	3.46 (0.16)	1.29 (0.25)	+174.68
Sells	$\hat{\mu}^{I*}$	$\hat{\mu}^{S*}$	$\hat{\sigma}^{I*}$	$\hat{\sigma}^{S*}$	\hat{c}_t^*	$1/\hat{\sigma}$	$\hat{\eta}^*$
	+0.93 (0.14)	+1.95 (0.66)	2.88 (0.10)	4.52 (0.96)	3.54 (0.17)	1.29 (0.25)	-179.28

Table 1 shows the median over all days of the point estimates per investor group G , in addition to the implied elasticity of demand ($\hat{\eta}$) and of supply ($\hat{\eta}^*$) on the platform, averaged across dealers. The corresponding medians of the standard errors are in parentheses. All estimates are in bps.

how many investors choose to trade bilaterally rather than on the platform. This is illustrated in Figure 4. It shows the distribution of yields that institutional buyers realize and a black line. Investors who draw liquidity shocks that would imply a bilateral yield that lies below the line buy on the platform, according to Proposition 1. Therefore, the position of the line—and, with it, the size of c_t —is determined by the fraction of investors who buy bilaterally rather than on the platform. Further, the shape of the yields’ distribution above the black line pins down the distribution of the liquidity shocks. This is because the investor realizes a yield $y_{tid}^I = \theta_t + \nu_{ti}^I - \hat{\xi}_{td}$ when buying bilaterally. Since we observe the trade yield (y_{tid}^I) and market value (θ_t) and we have already estimated dealer qualities ($\hat{\xi}_{td}$), we can solve for the liquidity shock (ν_{ti}^I) pointwise.

Finally, we back out the dealer’s value (v_{tj}^D) from the markup equation (7) of Proposition 2. We pick the v_{tj}^D for which the equation holds, given all the estimated parameters. This is similar to a classic approach adopted in industrial organization to infer the marginal costs of firms from firm behavior.

7 Estimation results

We report the estimates for a median day in Table 1—for example, $\hat{\mu}^I = \text{median}_t(\hat{\mu}_t^I)$. In all box plots, the upper and lower 1st percentile of the distribution is excluded.

Investor’s values. We find that buying retail investors are willing to pay about 2 bps more than institutional investors. When selling, the difference is smaller—about 1 bps—perhaps because retail investors who sell are more active than retail investors who only buy. This suggests that the yield gap of 4 bps between retail and institutional investors is not driven entirely by platform access, but that differences in the willingness to pay account for some of it. Below, we quantify how much.

Yield elasticity of demand. The yield elasticity of demand on the platform (which is driven mainly by the degree of platform competition) is about 174–179. This means that the demand of an institutional investor is relatively inelastic: Even if the dealer were willing to sell at a price at which she usually buys (which is about 0.5 bps higher), she would sell less than 1% more.

The elasticity of demand of an individual investor is similar, even though not directly comparable, to the aggregate elasticity of demand in the U.S. government bond market: Krishnamurthy and Vissing-Jørgensen (2012) estimate that the spread between corporate and government bond yields would increase by 1.5–4.25 bps if the debt/GDP ratio would rise by 2.5%. Our estimate implies a 2.6% increase in demand when the yield increases by 1.5 bps.

Dealer’s values and quality. Typically, a dealer values the bonds similar to an average institutional investor who is on the same side of the trade. This is plausible given that the average institutional investor is similar to a dealer, in that it is a large financial institution who frequently trades. However, there is large variation in the dealers’ values across days (see Figure 5). Further, while different dealers attach very similar (median) values to the bonds, they systematically differ in quality (see Figure 6b). This suggests that there might be welfare gains in matching investors to dealers who have higher values on that day or higher quality. We quantify these gains below.

Platform usage costs. In line with concerns that have been raised by industry experts, we find that high costs prevent investors from using the platform. At about 3.5 bps, the median cost is much larger than the actual fee to trade on the platform, which lies between 0.04 and 0.17 bps for a typical institutional investor.

Figure 5: Dealer values

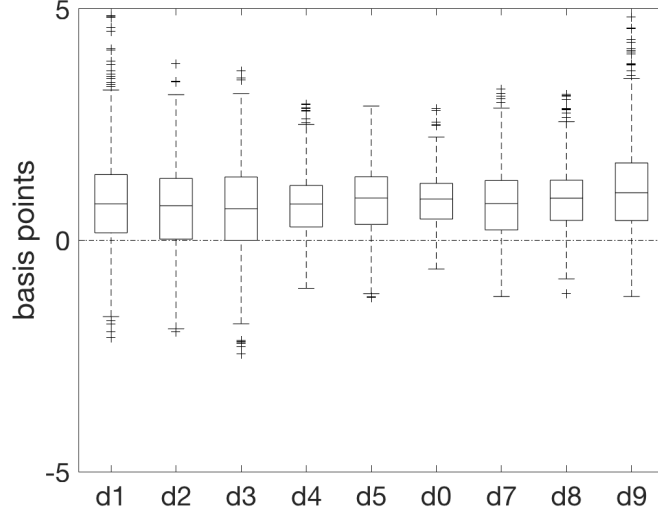


Figure 5 displays box plots of the estimated dealer values, \hat{v}_{tj}^D , net of the bond's market value, θ_t , in bps for each dealer. Dealers are labeled by d0 for the benchmark dealer to d8, and ordered according to their quality.

Figure 6: Dealer's quality

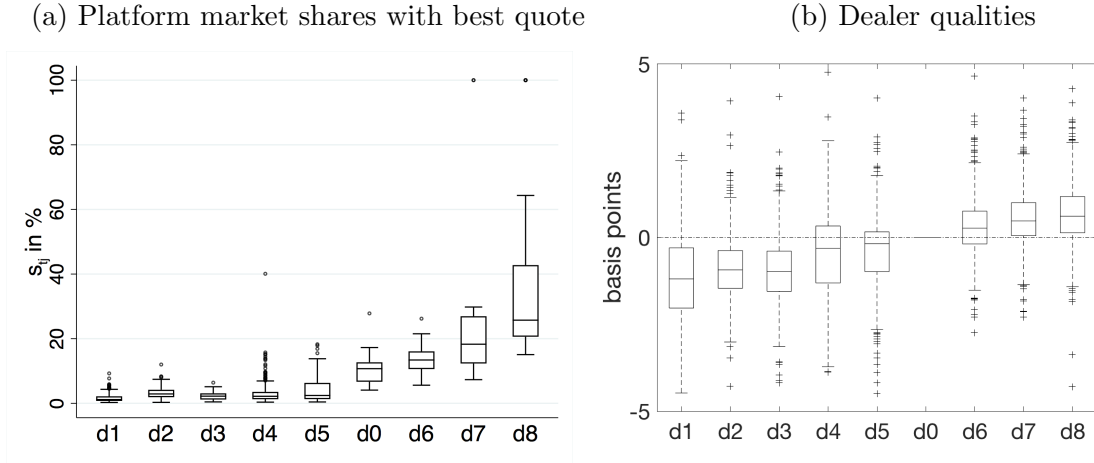


Figure 6a displays a box plot for each dealer, labeled d0 (benchmark) to d8. Each shows the distribution of how much of the total platform market this dealer captures (in %) on days on which she posts the best quote relative to other dealers. Figure 6b shows a box plot of the estimated qualities, $\hat{\xi}_{tj}$, in bps for each dealer.

Table 2: Model fit

	Event study		Model prediction	
Change in yield	-1.15	(0.340)	[-0.95, -0.98] for I	[-1.15, -1.24] for R

Table 2 compares by how much yields drop for investors who lose platform access with the standard error in parentheses (first column) with what the model predicts (second column). The former is the estimate of the event study regression but collapses the time before and after the event: $(y_{thsi j} - \theta_{ths})^+ = \zeta_i + \beta access_{thi} + \zeta_{th} + \zeta_s + \zeta_j + \epsilon_{thsi j}$, where $(y_{thsi j} - \theta_{ths})^+$ is defined as in (1) and $access_{thi}$ assumes value 1 if the investor has platform access and 0 otherwise. To compute the drop in the expected yield (before observing the liquidity shock) according to our model, we rely on Proposition 1. We keep the quotes constant at the observed levels, as in the event study.

7.1 Model fit

Before assessing the price and welfare effects from centralizing the market, we validate whether our parsimonious model can replicate the event study in Section 4. Recall that 90 institutional investors lost platform access in our sample. Crucially, we did not use any information on how yields change when this happens to estimate the model. Instead, we use this information to test whether our model predicts a similar impact on yields.

We find that the model’s prediction is very similar to the reduced-form estimate (see Table 2). On average, an institutional investor who loses platform access obtains a yield that is 1.15 bps worse in the data. Our model predicts that the yield drops by 0.95–1.24 bps. This similarity reassures us that the model makes adequate predictions about what happens when we make platform access universal in our counterfactual exercise.

8 Counterfactual exercises

We use the model to quantify how much of the gap between retail and institutional investors' yields is due to platform access and quantify welfare gains when further centralizing the market.

We do this by means of a counterfactual: We let retail and institutional investors have access to the platform on which dealers post quotes that are valid for any investor who uses the platform.²² We consider two specifications. In the first, all investors pay the estimated costs to use the platform. In the second, we set the usage costs to 0. This removes any type of friction that prevents investors from using the platform. If most of the usage costs are driven by privacy concerns, mandating anonymous trading might come close to this theoretical benchmark.

In all scenarios, we take into account how dealers and investors respond to the changes in the market rules: as investors enter the platform, dealers adjust their quotes, which in turn affects the trading decisions of investors. A new equilibrium arises. In this equilibrium, all investors select onto the platform, as in Proposition 1, and dealers set quotes similar to Proposition 2. Different from the status quo, however, dealers now behave as if there were a “representative” investor who draws liquidity shocks from a normal distribution with mean $\mu_t = \kappa^R \mu_t^R + \kappa^I \mu_t^I$ and standard deviation $\sigma_t = \kappa^R \sigma_t^R + \kappa^I \sigma_t^I$. Here, $\kappa^R = 0.1$ is the fraction of trades by retail investors and $\kappa^I = 0.9$ those by institutional investors on an average day.²³

Throughout, we take the ex ante perspective, which means that we take the expectation over how many retail versus institutional investors seek to trade and how much they are willing to pay. Further, we keep the number of trades fixed because our data do not allow us to estimate how likely it is that a trade is realized.

²²The findings are similar if we allow dealers to discriminate between investors on the platform and post a quote that is investor-group specific rather than a single quote for both groups.

²³We can show that this holds numerically, given our parameters. In theory, this is not always the case. The reason is that dealers do not take into account that retail investors may more strongly select onto the platform than institutional investors when setting quotes as if there were a representative investor.

For the welfare assessment, for example, this means that we focus on the question of who trades with whom and abstract from any gains or losses that may arise because more or fewer investors trade as market rules change.

8.1 What drives the yield gap?

When a retail investor obtains (costly) platform access, she expects a yield increase of about 1 bps. This implies that the gap between retail and institutional investors decreases on average by roughly 32% when the investor is buying. When the investor is selling, the percentage change is larger, at 47%, because the yield gap in the status quo is smaller.

The yield gap does not close completely, because many retail investors stay off the platform: Only 52%–60% of the retail investors would trade on the platform. The remaining would trade bilaterally. These investors obtain worse yields than institutional investors in bilateral trades because they are, on average, willing to pay more.

Platform participation is weak because it is costly to use the platform and because the platform is not perfectly competitive. To separate these two factors we eliminate the platform usage costs. Then, more retail investors (83%) would use the platform, and the yield gap would close by 52% for buying and 82% for selling investors. Some investors would still stay off the platform because of their low willingness to pay. For them, the platform quotes are not attractive.

8.2 Welfare analysis

Here we study how the total expected gains from trade, our welfare measure, changes. In Appendix D we analyze investor surpluses and dealer profits. We present all findings for investors who buy, but our findings generalize to selling investors, since the buy and sell sides of the market are close to symmetrical. Further, all findings are robust to assuming that the dealer only captures a fraction of—rather than the full—bilateral trade surplus (see Appendix E).

Definition 1. *The expected welfare is $W_t = \sum_G \kappa^G W_t^G$, where*

$$W_t^G = \sum_j \mathbb{E}[v_{tj}^D - v_{tj}^G(\boldsymbol{\nu}_{ti}^G) | \text{investor } i \in g \text{ buys from dealer } j] \quad (9)$$

is the expected welfare from trading with investors of group $G \in \{I, R\}$ with dealer value, v_{tj}^D , and investor value, $v_{tj}^G(\nu_{ti}^G) = \theta_t + \nu_{ti}^G - \xi_{tj}$. Proposition 1 specifies which dealer the investor buys from.

Whether welfare increases as more investors enter the platform depends on who matches with whom. To see this, we compute the change in welfare when going from the status quo to the counterfactual world:

$$\Delta W_t = \sum_G \kappa^G \sum_j \Delta \gamma_{ij}^G * (v_{tj}^D + \xi_{tj}). \quad (10)$$

Here, $\Delta \gamma_{ij}^G$ abbreviates the change in the probability that an investor in group G buys from dealer j on day t . Welfare increases as investors become more likely to buy from more efficient dealers, i.e., dealers with higher $(v_{tj}^D + \xi_{tj})$.²⁴

We find that welfare increases 9% when platform access becomes universal and by about 30% when access is free (see Figure 7). This translates into a sizable monetary gain of C\$ 123–411 million per year, or roughly 0.08–2.6 bps of GDP. The reason is that investors are more likely to match with more efficient dealers. For instance, a retail investor is 18% more likely to buy from the most efficient dealer.

To better understand where the welfare gain comes from, we decompose it into how much value is generated because dealers with higher values, v_{tj}^D , versus dealers with better quality, ξ_{tj} , are more likely to sell. We find that almost the entire welfare gain comes from matches to dealers with higher values.

This finding highlights the fact that in the status quo, dealers cannot freely sell

²⁴In theory, it is ambiguous whether welfare increases as more investors trade on the platform. On the platform, each investor i is free to choose among all of the dealers. She picks the dealer with the highest $(q_{tj} + \sigma \epsilon_{tij}) + \xi_{tj}$. This dealer must not be more efficient than the dealer who was chosen in the status quo, because the platform is not perfectly competitive. Dealers sell at quotes that differ from their values v_{tj}^D . They strategically set these quotes in response to investor behavior, and dealers with higher v_{tj}^D must not necessarily post higher quotes. Further, this negative effect is amplified when too many investors use the platform relative to what would be optimal if the platform were frictionless.

Figure 7: Welfare gain

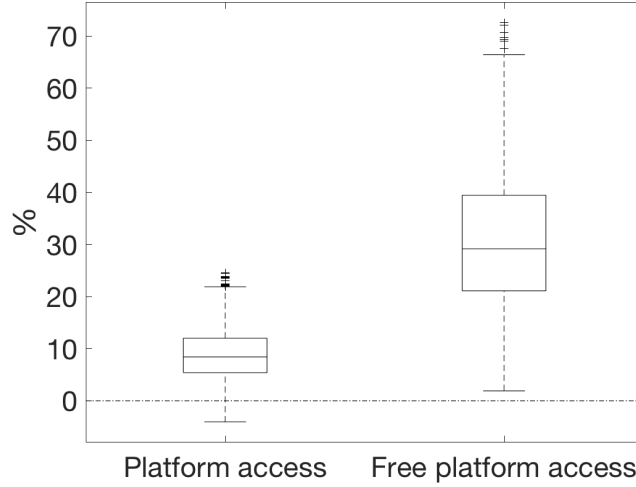


Figure 7 illustrates by how much welfare increases when making platform access universal or free. In both cases, it shows the distribution of the percentage change in welfare, $\Delta W_t/W_t * 100\%$, over days with W_t in (9), and ΔW_t in (10).

and buy as much they would like. For instance, a dealer who unexpectedly took a long inventory position might be more pressed to sell than a dealer who is short, but she might not be able to sell as much as she would like until the end of the day.

In March 2020, such frictions triggered dramatic events in the U.S. market for government bonds: When dealers failed to absorb enough bonds onto their balance sheets to meet the extraordinary supply of investors, the Federal Reserve System purchased trillions of U.S. government bonds and temporarily relaxed balance sheet constraints to rescue the market (Duffie (2020); He et al. (2020); Schrimpf et al. (2020)). Our findings suggest that market centralization would reduce these frictions.

Additional counterfactuals. So far, we have focused on market reforms that shift bilateral trading onto platforms on which investors run RFQ auctions with dealers. We view such a shift as a feasible first step in the right direction, but other reforms could affect trading. To assess the potential of other reforms to increase welfare, we quantify how efficient the market is today relative to the first best, in which a single

dealer—the one with the highest $v_{tj}^D + \xi_{tj}$ —sells to all investors on day t .²⁵

We compare four market settings to the first best: the status quo, the two counterfactuals in which all investors have access to the platform, and an additional counterfactual in which we remove the dealers and let investors directly trade with one another. The last counterfactual approximates an environment in which trades between investors realize via an efficient market mechanism, such as an efficient batch auction, as suggested by Budish et al. (2015): Each day, the market clears at the yield that equates expected investor demand with supply. All investors who seek to buy (sell) and are willing to accept a yield below (above) the market-clearing yield buy (sell). By assumption, dealers no longer participate in the market; for instance, because they no longer earn sufficient profits when the market clears via an efficient mechanism that minimizes markups.

Our findings are shown in Figure 8. The status quo achieves roughly 60% efficiency, which suggests that there are potentially large welfare gains from market reforms. Our first counterfactual, which allows all investors platform access at the estimated costs, does very little. The second counterfactual, in which we eliminate all costs, leads to a large increase in welfare, and we achieve 80% efficiency. Finally, letting investors trade directly with one another would lead to lower welfare than the status quo. This is because dealers no longer absorb the excess supply or demand of investors on days on which demand and supply do not balance perfectly. Crucially, this is not an endogenous outcome of running an efficient mechanism, but an assumption. Therefore, this finding highlights how important dealers are in providing liquidity, and should not be taken as an argument against efficient mechanisms.

Summary. Taken together, our findings suggest that even in a government bond market—which is commonly viewed as one of the most well-functioning financial markets—there is large potential to increase welfare by centralizing the market.

²⁵This is because, in our model, over the course of a day dealers have constant values. In turbulent times, this might no longer be the case and first best would be different.

Figure 8: Market efficiency

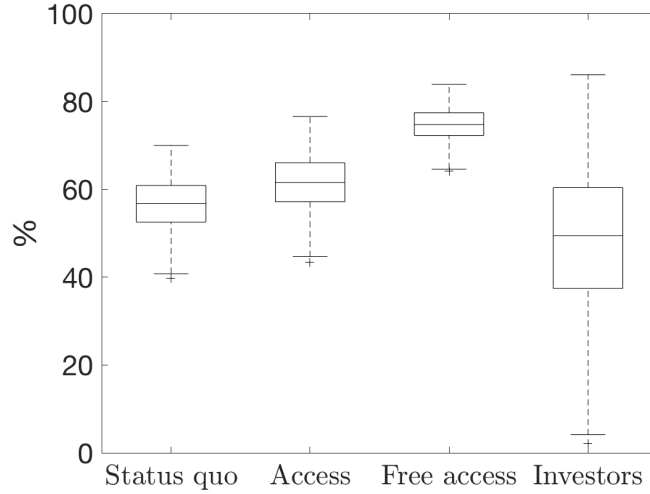


Figure 8 shows the distributions of daily welfare W_t , defined in (9), as the percentage of what could be achieved in the first best in four settings: the status quo, the counterfactuals in which all investors have platform access at the estimated platform usage costs and for free, and the counterfactual in which investors trade with one another.

8.3 Robustness

We conduct several tests to verify the robustness of our findings in Appendix E. First, we test the robustness of our parameter estimates. For example, we check whether the estimates are biased in the expected direction when we do not instrument the quotes or use the amount a dealer won as instrument. We also verify that measurement errors in the dealers' qualities (ξ_{tj}) do not significantly bias the distribution of the liquidity shocks. In addition, we allow for dealer-specific platform usage costs (c_{tj}) and restrict the sample to exclude occasionally large trades.

Second, we verify that our estimates and welfare findings are robust when we allow the investor to capture some trade surplus in the bilateral trade. For this, we rely on the extended model in Appendix B.2, in which the investor captures a trade surplus of ϕ in a bilateral trade. While we cannot identify this parameter with our data, we can test how the model estimates and counterfactual findings change as we increase ϕ from zero (as in the benchmark model) to positive values.

Taken together, our robustness tests confirm our expectations and suggest that our main findings are qualitatively robust.

9 Conclusion

In this paper, we use trade-level data on the Canadian government bond market to study whether to centralize OTC markets by shifting bilateral trades onto multi-dealer platforms on which dealers compete for investors. We show that even in a seemingly frictionless market, platform access can lead to better prices for investors. Further, we estimate large welfare gains because more trades are intermediated by dealers who urgently seek to trade. We expect this to be true for many other OTC markets.

References

- Abudy, M. M. and Wohl, A. (2018). Corporate bond trading on a limit order book exchange. *Review of Finance*, pages 1413–1440.
- Allen, J., Kastl, J., and Wittwer, M. (2020). Primary dealers and the demand for government debt. Working paper.
- Babus, A. and Parlatore, C. (2019). Strategic fragmented markets. Working paper.
- Baldauf, M. and Mollner, J. (2020). Principal trading procurement: Competition and information leakage. Working paper.
- Barklay, M. J., Hendershott, T., and Kotz, K. (2006). Automation versus intermediation: Evidence from Treasuries going off the run. *Journal of Finance*, 66(5):2395–2414.
- Benos, E., Payne, R., and Vasios, M. (2020). Centralized trading, transparency, and interest rate swap market liquidity: Evidence from the implementation of the Dodd-Frank Act. *Journal of Financial and Quantitative Analysis*, 55(1):159–192.
- Bessembinder, H., Spatt, C., and Venkataraman, K. (2020). A survey of the microstructure of fixed-income markets. *Journal of Financial and Quantitative Analysis*, 55(1):1–45.

- Biais, B. and Green, R. (2019). The microstructure of the bond market in the 20th century. *Review of Economic Dynamics*, 33:250–271.
- Boyarchenko, N., Lucca, D., and Veldkamp, L. (2021). Taking orders and taking notes: Dealer information sharing in treasury auctions. *Journal of Political Economy*, forthcoming.
- Bresnahan, T. F. (1981). Departures from marginal-cost pricing in the American automobile industry: Estimates for 1977–1978. *Journal of Econometrics*, 17(2):201–227.
- Budish, E., Cramton, P., and Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics*, 130(4):1547–1621.
- Chen, D. and Duffie, D. (2021). Market fragmentation. *American Economic Review*, forthcoming.
- De Roure, C., Moench, E., Pelisson, L., and Schneider, M. (2020). OTC discount. Working paper.
- Duffie, D. (2020). Still the world’s safe haven? Redesigning the U.S. Treasury market after the COVID-19 crisis. Working paper.
- Duffie, D., Dworczak, P., and Zhu, H. (2017). Benchmarks in search markets. *Journal of Finance*, 72(5):1983–2044.
- Duffie, D., Gârleanu, N., and Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, 73(6):1815–1847.
- Dugast, J., Üslü, S., and Weill, P.-O. (2019). A theory of participation in OTC and centralized markets. Working paper.
- Dunne, P. G., Hau, H., and Moor, M. J. (2015). Dealer intermediation between markets. *Journal of the European Economic Association*, 13(5):770–804.
- Financial Times* (2015). Wall street banks to settle CDS lawsuit for \$1.9bn. Ben McLannahan and Joe Rennison.
- Fleming, M. J., Mizrach, B., and Nguyen, G. (2017). The microstructure of a U.S. Treasury ECN: The BrokerTec platform. Federal Reserve Bank of New York Staff Reports.

- Friewald, N. and Nagler, F. (2019). Over-the-counter market frictions and yield spread changes. *Journal of Finance*, 74(6):3217–3257.
- Garbade, K. D. and Silber, W. L. (1976). Price dispersion in the government securities market. *Journal of Political Economy*, 84(4):721–740.
- Glode, V. and Opp, C. C. (2019). Over-the-counter versus limit-order markets: The role of traders’ expertise. *Review of Financial Studies*, 33(2):866–915.
- Green, R. C., Hollifield, B., and Schürhoff, N. (2007). Dealer intermediation and price behavior in the aftermarket for new bond issues. *Journal of Financial Economics*, 86:643–682.
- Hau, H., Hoffmann, P., Langfield, S., and Timmer, Y. (2019). Discriminatory pricing of over-the-counter derivatives. Working paper.
- He, Z., Nagel, S., and Song, Z. (2020). Treasury inconvenience yields during the COVID-19 crisis. Working paper.
- Hendershott, T., Li, D., Livdan, D., and Schürhoff, N. (2020a). Relationship trading in OTC markets. *Journal of Finance*, 75(2):683–734.
- Hendershott, T., Li, D., Livdan, D., and Schürhoff, N. (2020b). True cost of immediacy. Working paper.
- Hendershott, T. and Madhavan, A. (2015). Click or call? Auction versus search in the over-the-counter market. *Journal of Finance*, 80(1):419 – 447.
- Hortaçsu, A. and Kastl, J. (2012). Valuing dealers’ informational advantage: A study of Canadian Treasury auctions. *Econometrica*, 80(6):2511–2542.
- Hortaçsu, A. and McAdams, D. (2010). Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish Treasury auction market. *Journal of Political Economy*, 118(5):833–865.
- Kastl, J. (2011). Discrete bids and empirical inference in divisible good auctions. *Review of Economic Studies*, 78:974–1014.
- Koijen, R. S. and Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515.
- Koijen, R. S. and Yogo, M. (2020). Exchange rates and asset prices in a global demand system. Working paper.

- Kondor, P. and Pintér, G. (2019). Clients’ connections: Measuring the role of private information in decentralized markets. Working paper.
- Krishnamurthy, A. and Vissing-Jørgensen, A. (2012). The aggregate demand for Treasury debt. *Journal of Political Economy*, 120(2):233–267.
- Krishnamurthy, A. and Vissing-Jørgensen, A. (2015). The impact of Treasury supply on financial sector lending and stability. *Journal of Financial Economics*, 118:571–600.
- Liu, Y., Vogel, S., and Zhang, Y. (2018). Electronic trading in OTC markets vs. centralized exchange. Swiss Finance Institute Research Paper No. 18-19.
- Logan, L. K. (2020). Speech: Treasury market liquidity and early lessons from the pandemic shock. Accessed on 05/05/2021 at: <https://www.newyorkfed.org/newsevents/speeches/2020/log201023>.
- Loon, Y. C. and Zhong, Z. K. (2016). Does Dodd-Frank affect OTC transaction costs and liquidity? Evidence from real-time CDS trade reports. *Journal of Financial Economics*, 119:645–672.
- Managed Funds Association (2015). Why eliminating post-trade name disclosure will improve the swaps market. MFA position paper.
- McPartland, K. (2016). Corporate bond electronic trading continues growth trend. Industry report.
- Monias, S., Nguyen, M., and Valentex, G. (2017). Funding constraints and market illiquidity in the European Treasury bond market. Working paper.
- O’Hara, M. and Zhou, X. (2021). The electronic evolution of corporate bond dealers. *Journal of Financial Economics*, forthcoming.
- Plante, S. (2017). Should corporate bond trading be centralized? Working paper.
- Riggs, L., Onur, E., Reiffen, D., and Zhu, H. (2020). Swap trading after Dodd-Frank: Evidence from index CDS. *Journal of Financial Economics*, 137:857–886.
- Rostek, M. J. and Yoon, J. H. (2020). Exchange design and efficiency. Working paper.
- Schrimpf, A., Shin, H. S., and Sushko, V. (2020). Leverage and margin spirals in fixed income markets during the COVID-19 crisis. BIS Bulletin.
- Vogel, S. (2019). When to introduce electronic trading platforms in over-the-counter markets? Working paper.

- Wittwer, M. (2020). Interconnected pay-as-bid auctions. *Games and Economic Behavior*, 121:506–530.
- Wittwer, M. (2021). Connecting disconnected financial markets? *American Economic Journal: Microeconomics*, 13(1):252–282.

Online Appendix

A Data cleaning

For a few trades of the 3,755,901 observations in the raw data, we change the execution time, the date, or the settlement date. First, 296 trades were reported on a weekend. We count them as Monday 7:00 am trades if reported on a Sunday and Friday 5:00 pm trades if reported on a Saturday. In all other cases, we keep the time of day and only change the date. 162 observations are reported to settle after maturity. We replace their settlement date with the maturity date. 5,355 trades settle before they were executed. We replace the reported settlement date with the date on which the trade would settle according to trade conventions.

We correct 100 cases in which subsidiaries of reporting dealers or brokers are labeled retail investors, and we drop 20 observations that were reported without retail/institutional indicator. Of the investors who switch from retail to institutional or vice versa, we drop investors who do not permanently switch. This excludes trades with investors who are in a gray area. For example, CIBC Investor Services Inc., a subsidiary of the Canadian Imperial Bank of Commerce (CIBC), classifies as a retail investor according to the rules, even though CIBC is one of biggest banks in Canada. A reporting dealer who trades with CIBC Investor Services Inc. might falsely believe that this investor is institutional and report it as such.

We exclude trades that exhibit yields that are extreme relative to the public Bloomberg mid-yield since it is difficult to rationalize why anyone would be willing to accept these trades. They could be reporting errors or part of a larger investment package that we do not observe. To detect these outliers, we analyze the distribution of the markup $(y_{thsi} - \theta_{ths})^+$, defined in (1). It has extremely long but very thin tails. We drop the upper and lower 1% of this distribution for each investor group.

We focus on CanDeal or bilateral trades only, which means that we ignore 0.41% of the observations with incorrect trading venues. In these rare cases, the dealer makes a mistake and typically reports the ID of her counterparty as the trading venue.

Finally, in rare cases in which a Bloomberg quote for a security is missing in an hour of the day, we use the daily average Bloomberg quote of this security.

B Mathematical appendix

B.1 Proof of Propositions 1 and 2

The equilibrium can be derived by backward induction. For notational convenience, we drop the subscript t and the superscript I throughout the proof.

Proposition 1. Statement (i) holds by assumption. To derive statement (ii), begin in the last stage. Conditional on entering the platform and observing ϵ_{ij} , investor i buys from dealer j if $u_{ij} > u_{ki} \forall k \neq j$, where $\tilde{u}_{ij}(\epsilon_{ij}) = \xi_t + q_j + \sigma \epsilon_{ij}$. Ex ante, dealer j 's market share on the platform (of investors on the platform) is

$$s_j(q) = \frac{\exp(\delta_j)}{\sum_k \exp(\delta_k)} \text{ with } \delta_j = \frac{1}{\sigma}(\xi_j + q_j) \text{ given } \epsilon_{ij} \sim EV1. \quad (11)$$

By assumption, a home dealer d offers $y_{id} = \theta + \nu_i - \xi_d$ in a bilateral trade and is always willing to trade. The investor obtains no surplus when buying bilaterally and expects to earn $-(\theta + \nu_i) + \mathbb{E}[\max_{k \in \mathcal{J}} \tilde{u}_{ki}(\epsilon_{ki})] - c$ when entering the platform. She decides to buy bilaterally if $\psi(q) \leq \nu_i$ with $\psi(q) = \mathbb{E}[\max_{k \in \mathcal{J}} \tilde{u}_{ki}(\epsilon_{ki})] - \theta - c$. \square

Proposition 2. Consider home dealer d . In choosing the quote, the dealer anticipates how investors will react but does not know which liquidity shocks investors will draw. The dealer chooses q_d to

$$\max_{q_d} \pi_d(q) = \max_{q_d} \{\pi_d^D(q) + \pi_d^E(q)\}, \text{ where}$$

$$\pi_d^D(q) = \int_{\psi(q)}^{\infty} (value_d - (\theta + \nu - \xi_d)) f(\nu) d\nu \text{ given that } y_d = \theta + \nu - \xi_d$$

is the expected profit from bilateral trades, and

$$\pi_d^E(q) = S_d(q)(value_d - q_d), \text{ where } S_d(q) = \sum_j F(\psi(q)) * s_d(q) \text{ and } s_d(q), \text{ as in (11),}$$

is the expected profit from platform trades. Taking the partial derivative w.r.t. q_d , and rearranging gives the markup equation. \square

B.2 Model extension: Bargaining power

Let the dealer and investor bargain under complete information and denote the investor's bargaining power by $\phi \in [0, 1]$. The dealer and an investor of group G agree on the following bilateral yield: $y_{tid} = \phi v_{tj}^D + (1 - \phi)v_{tij}^G$, where v_{tj}^D and v_{tij}^G are the dealer's and investor's value, respectively. The equilibrium characterization can be derived analogous to the benchmark model and is omitted here.

B.3 Micro-foundation

The goal of this section is to micro-found why the yield at which investor i trades with dealer j on the platform trades is $q_{tj} + \sigma_{\epsilon_{tij}}$ on day t . For simplicity, let dealers be ex ante identical—i.e., abstract from dealer qualities. Later on, we explain how one could extend the theoretical model to include the quality term. Further, we restrict attention to institutional investors who have access to the platform. Including retail investors with no access is straightforward and analogous to the structural model.

Dealers. $J \geq 2$ dealers sell a bond to investors. Each transaction is a single unit trade. Each investor has a home dealer, called d , short for d_i , and each dealer has a home investor base. It consists of a unit mass of investors.

A dealer aims at maximizing profit. Ex post, the dealer obtains a profit of $v - y$, when selling one unit at yield y and valuing the bond by v . The dealer's value splits into a part that is commonly known to all dealers, v_1 , and a part that is unknown, v_2 , which is drawn iid from a commonly known normal distribution:

$$v = v_1 + v_2 \text{ with } v_1 \in \mathbb{R}^+ \text{ and } v_2 \stackrel{iid}{\sim} N(\mu_v, \sigma_v^2).$$

Investors. All investors can either buy bilaterally from their home dealer or use the platform. In making this decision, each investor i aims at maximizing surplus. She obtains a surplus of $y - \nu_i$ when buying from a dealer at yield y , and valuing the bond by

$$\nu_i \stackrel{iid}{\sim} N(\theta + \mu_\nu, \sigma_\nu^2), \text{ where } \theta \in \mathbb{R}.$$

Both θ and the normal distribution are commonly known.

Timing of events. The game has two stages. In the first stage, each dealer j posts a quote, q_j , on the platform simultaneously with all other dealers. The quote signals the benchmark yield an investor can obtain when buying from this dealer on the platform. In the second stage, each investor observes her private value, ν_i , and decides between buying bilaterally from her home dealer or on the platform. In a bilateral trade, the home dealer observes ν_i and charges $y_i = \nu_i$. On the platform, the investor runs a first-price auction with all dealers.²⁶ To do so, she has to pay a cost $c \in \mathbb{R}^+$.

Before running an auction with investor i , each dealer j draws a private signal x_{ij} about the common value component of her value:

$$x_{ij} = v_2 + s\omega'_{ij}, \text{ where } \omega'_{ij} \stackrel{iid}{\sim} N(0, 1) \text{ and } s \in \mathbb{R}^+.$$

Given these signals, each dealer submits her bid. In choosing the bid, dealers want to avoid the posted quotes being seen as cheap talk. They agree that each dealer should be allowed to offer a bid that differs from the posted quote for a given realization of the signal, but that must be proportional to the quote in expectation. The easiest way to achieve this outcome is to collude and bid as if their value v_1 were the posted quote. This is profitable, because in equilibrium the posted quotes are smaller than v_1 .²⁷

Equilibrium. In a symmetric equilibrium, dealers and investors behave similarly to how they behave in the structural model. One can characterize equilibrium conditions that are analogous to those in Propositions 1 and 2. Here, we focus only on the characterization of the equilibrium bids.

²⁶We could also assume that the investor randomly picks a subset of dealers.

²⁷Without imposing such collusion, the platform yield has a similar functional form to (12) of Proposition 3, stated below. There would be an additional term that depends on the equilibrium quote and the elasticities that govern how much lower this quote is relative to the dealer's value, v_1 . If this were the true data-generating process, our estimates would be a close approximation of reality if this additional term were small enough.

Proposition 3. *In a symmetric equilibrium, in which all dealers post the same quote q^* , the dealer with the highest signal, x_{ij} , wins the auction, and the investor obtains the following yield on the platform:*

$$y_{ij}^E = q^* + \sigma \epsilon_{ij} \text{ where } \epsilon_{ij} = (x_{ij}/\sigma + s) \text{ and} \quad (12)$$

$$\sigma \text{ solves } 0 = \int_{-\infty}^{\infty} [-\Phi(-z)^{J-1} + (z - \sigma)(J - 1)\Phi(-z)^{J-2}\phi(-z)] \phi(z) dz. \quad (13)$$

$z \sim N(0, 1)$ and $\Phi(\cdot)$, $\phi(\cdot)$ are the CDF, PDF of the standard normal distribution.

Proof. We derive conditions that are satisfied in a symmetric equilibrium via backward induction. For notational convenience, we drop subscripts whenever possible.

Guess that there is an equilibrium in which a dealer with signal x submits a linear function $\beta(x) = q^* + x + \sigma s$, where σ is a parameter, and s determines the noisiness of the dealer's signal.

Note that conditional on x , $v_2|x \sim N([\mu_v/h + x/h']/[h + h'], 1/[h + h'])$, where $h = 1/\sigma_v^2$, $h' = 1/s^2$. As $\sigma_v^2 \rightarrow \infty$, $h \rightarrow 0$, and at limit $v|x \sim N(x, s^2)$. Hence $[v_2 - x]/s = \omega$ where $\omega \sim N(0, 1)$.

The dealer's problem is, given x ,

$$\arg \max_b \int_{-\infty}^{\infty} [q^* + v_2 - b] F(\beta^{-1}(b)|v_2)^{J-1} dF(v_2|x),$$

where $F(\beta^{-1}(b)|v_2)^{J-1}$ is the distribution of maximum of others' signals given v_2 , evaluated at $\beta^{-1}(b)$, and $F(v_2|x)$ is the distribution of v_2 given signal x . The first-order condition is

$$0 = \int_{-\infty}^{\infty} [-F(z|v_2)^{J-1} + (q^* + v_2 - \beta(x))(J - 1)F(z|v_2)^{J-2}f(z|v_2)\frac{\beta^{-1}(b)}{db}] dF(v_2|x),$$

evaluated at $z = \beta^{-1}(\beta(x)) = x$, so $\frac{d\beta^{-1}(\beta(x))}{db} = 1/\beta'(x)$.

To evaluate $F(x|v_2)$, observe that it is the probability that another bidder's signal $x' < x$ given v_2 . Since $x' = v_2 + \omega's$ this is the probability that $v_2 + \omega's < x$ or $\omega' < [x - v_2]/s$, which is $F(x|v_2) = \Phi([x - v_2]/s)$. From above, $[x - v_2]/s = -\omega$, so $F(x|v_2) = \Phi(-\omega)$ and $f(x|v_2) = \phi(-\omega)/s$. Lastly, since $[v_2 - x]/s = \omega$, $F(v_2|x) =$

$\Phi(\omega)$ and $f(v_2|x) = \phi(\sigma)/s$. Taken together, the first-order condition becomes

$$0 = \int_{-\infty}^{\infty} [-\Phi(-\omega)^{J-1}\beta'(x) + (q^* + x + \omega s - \beta(x))(J-1)\Phi(-\omega)^{J-2}\phi(-\omega)/s] \phi(\omega)/s d\omega.$$

Given $\beta(x) = q^* + x + \sigma s$, the FOC is satisfied when σ solves (13). \square

Like in the structural model, on the platform the investor receives a yield that equals the posted quote plus a stochastic term. There are two differences.

First, in the structural model, dealers differ in quality, ξ_j , and therefore post different quotes. We could extend the model and include a quality term ξ_j in the investor's surplus. In this case, dealers would post different quotes depending on their quality. To keep the model tractable, one would have to assume that the auction determines the markup (or discount) over the posted quote and that the dealer with the best markup wins, independent of the posted quote. Otherwise, the auction would become asymmetrical.

Second, the platform shocks are iid in the structural model but may be correlated here. The iid assumption is common in the literature on demand estimation, but, as the micro-foundation highlights, imposes a restriction. To achieve independence in the theoretical model, one would have to assume that the dealers draw independent signals, conditional on their value v_1 .

C Details regarding estimation

C.1 Construction of the supply shock instruments

In modeling the auction and estimating the bidders' values, we follow Hortaçsu and Kastl (2012) and Allen et al. (2020). For a detailed discussion of all assumptions and derivation of the equilibrium, we refer to these papers.

Auction model. In the auction, there are two groups of bidders: N_d dealers and N_c customers, who are investors who bid at auction. All of them draw a private signal.

Assumption 5. *Dealers' and customers' private signals s_j^d and s_j^c are for all bidders j independently drawn from common atomless distribution functions F^d and F^c with support $[0, 1]^M$ and strictly positive densities f^d and f^c .*

The bidder's group and signal affect how much she values the bond.

Assumption 6. *A bidder j of group $g \in \{d, c\}$ with signal s_j^g values amount q by $v^g(q, s_j^g)$. This value function is nonnegative, measurable, bounded strictly increasing in s_j^d for all q and weakly decreasing in q for all s_j^g .*

Given their values, bidders place bids. Each bid is a step function that characterizes the price the bidder would like to pay for each amount.

Assumption 7. *Each bidder has the following action set:*

$$A = \left\{ \begin{array}{l} (b, q, K) : \dim(b) = \dim(q) = K \in \{1, \dots, \overline{K}\} \\ b_k \in [0, \infty) \text{ and } q_k \in [0, 1] \\ b_k > b_{k+1} \text{ and } q_k > q_{k+1} \forall k < K. \end{array} \right.$$

Dealers can submit their bids directly to the auctioneer (the Bank of Canada). Customers have to place their bids with one of the dealers. This might give the dealer additional information. To capture this, we define the information that is available to dealer j before placing her bid by Z_j . We call $\theta_j^d = (s_j^d, Z_j)$ the dealer's type. The type of a customer is her private signal s_j^c .

Definition 2. *A pure strategy is a mapping from the bidder's set of types to the action space: $\Theta_j^g \rightarrow A$. It is a bidding function, labeled $b_j^g(\cdot, \theta_j^g)$ for bidder j of group g with type θ_j^g .*

Once all bidders submit their step function, the market clears at the lowest price at which the aggregated submitted demand satisfies the total supply.

The supply is unknown to each bidder when she places her bid because a fraction of it goes to noncompetitive tenders. These are bids that specify only an amount that is won with certainty.

Assumption 8. Supply \mathbf{Q} is a random variable distributed on $[\underline{Q}, \overline{Q}]$ with strictly positive marginal density conditional on $s_j^g \forall i, g = c, d$.

Bidder j wins amount q_j^c at market clearing and pays how much she offered to win for each unit she won.

Definition 3. A Bayesian Nash equilibrium in pure strategies is a collection of functions $b_j^g(\cdot, \theta_j^g)$ that for each bidder j and almost every type θ_j^g maximizes the expected total surplus, $\mathbb{E} \left[\int_0^{q_j^c} [v(x, s_j^g) - b_j^g(x, \theta_j^g)] dx \right]$.

We focus on type-symmetric BNE in which all dealers and all customers play the same strategy. One can show that in any type-symmetric BNE, every step k in the bid function $b_j^g(\cdot, \theta_j^g)$ has to satisfy

$$v^g(q, s_j^g) = b_k + \frac{\Pr(b_{k+1} \geq \mathbf{P}^c | \theta_j^g)}{\Pr(b_k > \mathbf{P}^c > b_{k+1} | \theta_j^g)} (b_k - b_{k+1}) \quad (14)$$

for all but the last step and $b_k = v^g(\bar{q}(\theta_j^g), s_j^g)$ at the last step, where $\bar{q}(\theta_j^g)$ is the maximal amount the bidder may be allocated in equilibrium. Here P^c denotes the market-clearing price.

Estimation. We estimate how much each dealer expects to win in this equilibrium, at the time at which she places the bids, in four steps.

First, we estimate the distribution of the residual supply curve a dealer faces. This curve is the total supply minus the total demand of all other bidders. For this, we draw N_c customer bids from the empirical distribution of customer bids in the auction, replacing bids by customers who did not bid in the auction with 0. We then find the dealer(s) who observed each of the customer bids and draw their bids. In rare cases, in which the customer submitted more than one bid, we draw bids uniformly from all dealers who observed this customer. If at that point the total number of dealers we have already drawn is still lower than the number of potential dealers minus one, we draw the remaining dealer bids from the pool of dealers who do not observe a customer bid.

We then let the market clear for each realization of the residual supply curve. This gives the distribution of how much the dealer won in the auction, q_j^c . It also specifies,

for each step of the dealer's bidding function, how likely it is that the market clears at that step, i.e., that $b_k \geq \mathbf{P}^c > b_{k+1}$.

With that, we can compute how much the dealer expected to win when bidding:

$$\mathbb{E}[\text{amount dealer } j \text{ wins} | \text{bids}] = \sum_k^{K_j} \hat{\Pr}(b_k \geq \mathbf{P}^c > b_{k+1} | \theta_j^d) * \hat{\mathbb{E}}[\mathbf{q}_j^c | b_k \geq \mathbf{P}^c > b_{k+1}, \theta_j^d],$$

where K_j are the steps in dealer j 's bidding function.

Finally, to obtain our instrument $won_{j\tilde{t}}$ of auction \tilde{t} , we subtract $\mathbb{E}[\text{amount dealer } j \text{ wins} | \text{bids}]$ from the amount that bidder j actually won, which we observe.

C.2 Main estimation procedure

We explain the estimation for buying institutional investors in detail. For buying retail investors, we match the expectation and variance of the bilateral yields via GMM (similar to step 2 below). For selling investors, the estimation is analogous.

Step 1. Denote dealer j 's market share on the platform by $s_{tj}(q_t, \xi_t, \sigma)$. When ϵ_{tij} are extreme value type 1 distributed:

$$s_{tj}(q_t, \xi_t, \sigma) = \frac{\exp(\delta_{tj})}{\sum_{k \in \mathcal{J}_t} \exp(\delta_{tk})} \text{ with } \delta_{tk} = \frac{1}{\sigma}(\xi_{tk} + q_{tk}) \text{ for all } k \in \mathcal{J}_t. \quad (15)$$

Abbreviate $s_{tj}(q_t, \xi_t, \sigma)$ by s_{tj} for all j , divide this expression for all $j \neq 0$ by the equivalent expression for the benchmark dealer ($j = 0$), take logs, and use Assumption 2 plus Normalization 1 to obtain

$$\log(s_{tj}/s_{t0}) = \zeta_j + \zeta_t + \frac{1}{\sigma} \tilde{q}_{tj} + rest_{tj}, \quad (16)$$

where $\tilde{q}_{tj} = q_{tj} - q_{t0}$, $\zeta_j = \frac{1}{\sigma} \xi_j$, $\zeta_t = \text{mean}_j \left(\frac{1}{\sigma} \chi_{tj} \right)$, $rest_{tj} = \frac{1}{\sigma} \chi_{tj} - \zeta_t$.

Under Assumption 3, which implies $\mathbb{E}[\mathbf{rest}_{tj} | won_{\tilde{t}j}, \zeta_t, \zeta_j] = 0$, we can estimate σ in a linear IV regression in which we instrument \tilde{q}_{tj} by $won_{\tilde{t}j}$ and include dealer ζ_j and date ζ_t fixed effects.

With this, we compute $\hat{\xi}_{tj}$ for all $j \neq 0$ and the cutoff that determines whether an investor buys bilaterally from home dealer d or on the platform:

$$\mathbb{E}[\max_{k \in \mathcal{J}_t} \tilde{u}_{tik}(\boldsymbol{\epsilon}_{t\mathbf{i}k})] = \hat{\sigma} \ln \left(\sum_{k \in \mathcal{J}_t} \exp \left(\frac{1}{\hat{\sigma}} (q_{tk} + \hat{\xi}_{tk}) \right) \right).$$

Step 2. For each day t , we estimate the remaining parameters via GMM by matching the expectation and variance of the bilateral yield of a buying institutional investor, and the probability that such an investor buys bilaterally. To compute the predicted moments, we rely on $\boldsymbol{\nu}_{\mathbf{i}}^I \sim N(\mu_t^I, \sigma_t^I)$ and Proposition 1.

D Additional findings: Who wins and who loses?

To assess who wins and loses, we compute by how much the expected investor surplus and dealer profit (prior to observing the liquidity and platform shocks) change when going from the status quo to the counterfactual market rules.

In theory, retail investors cannot lose surplus when obtaining platform access, but it is unclear whether institutional investors or dealers benefit from the change. This depends on how dealers adjust their quotes. If they set quotes that are more favorable to investors, dealers lose and institutional investors win.

Whether this is the case depends on how the two elasticities that govern the quotes change as the composition of investors on the platform changes. A lower elasticity of demand on the platform toughens platform competition and therefore leads to better quotes. A lower cross-market elasticity has the opposite effect: The less easily investors switch onto the platform, the lower the incentive for dealers to post unattractive quotes to prevent investors from using the platform. Which of the two effects dominates is an empirical question.

We find that the second effect dominates slightly, so that quotes become better for investors as platform access becomes universal. When allowing retail investors to enter the platform at estimated costs, the quotes change very little. The reason is that quotes are more strongly targeted to institutional investors, who make up 90% of the market, and institutional investors already have platform access in the status quo. However, as we eliminate platform usage costs, the average quote increases by about 0.05 bps, which is roughly 1/10 of the median bid-ask spread.

As a result, free platform access brings higher gains to investors and larger losses to

dealers than costly platform access. Retail investors gain about 4 bps and institutional investors about 1 bps, and dealers lose about 1 bps (per unit) when access is free.

For an average retail investor, who trades about C\$ 86 million (units) per year, this is equivalent to earning about $4\text{bps} \times \text{C\$ } 86 = \text{C\$ } 34$ thousand more interest in a year. For an average institutional investor, the monetary gain is larger because institutional investors trade larger amounts and more often in a year than retail investors: C\$ 1.7 million. This gain is significant; it would double the revenue an average institutional investor makes from interest on any type of investment in a year. Dealers, who trade the most, each expect a monetary loss of about C\$ 27 million per year. This is a sizable reduction in the dealer’s yearly revenue from interest on any type of investment of 15%.

E Robustness analysis

The first set of robustness checks mainly concerns the estimate of the competition parameter, σ , and the dealer’s qualities, ξ_{tj} (see Appendix Table 7).

First, we check whether $\hat{\sigma}$, which governs the yield elasticity of demand on the platform, is biased in the expected direction when we do not instrument the quotes by won_{tj} and replace Assumption 3 with $\mathbb{E}[\chi_{tj}|\zeta_t, \xi_j] = 0$. The OLS estimate implies an elasticity that is close to zero. The endogeneity bias goes in the expected direction. It comes from a misspecified estimate of $\sigma > 0$, which is biased downward if dealers decrease the yield quote (i.e., increase the price) in response to higher demand for reasons that are unobservable to the econometrician.

Next, we use a different instrument for the quote; namely, the amount a dealer won on the most recent auction day rather than the amount she won unexpectedly. The advantage of this instrument is that it is model-free because we can read it off the data. The big disadvantage is that it does not address the concern that dealers might anticipate investor demand and bid accordingly in the auction.

We obtain similar estimates with both instruments when allowing for dealers to systematically differ in quality. However, the instrument is weak. We therefore check by how much the estimates change when dropping the dealer fixed effect by imposing

$\xi_j = 0 \forall j$. In both specifications, this increases the correlation between our instrument and the platform quotes. The instrument becomes stronger, but we no longer control for unobservable differences between dealers that might drive differences in the quotes. $\hat{\sigma}$ decreases from 0.77 to 0.49 when using $won_{\bar{t}j}$ and from 0.68 to 0.62 when using the amount the dealer actually won as instrument.

The second set of robustness tests regards the estimates of the platform usage costs and the distribution of the liquidity shocks. To verify that the distribution of the liquidity shocks is not much biased by measurement errors in the quality of the dealers, ξ_{tj} , we estimate the model under the assumption that the bilateral yield equals the market value plus the liquidity shock, i.e., $y_{ti}^G = \theta_t + \nu_{ti}^G$ (see Appendix Table 8).

Further, we check by how much the estimates change when we allow the platform usage cost to be dealer specific, c_{tj} . The dealer-specific cost is identified from how many trades each dealer j realizes on versus off the platform in period t rather than how many trades are realized on versus off the platform in total. To obtain sufficient power for each dealer, we pool five trading days and let the period t be a business week, rather than a single day. This implies that we cannot directly compare the point estimates in Table 9 with the estimates of our main specification in Appendix Table 1. For instance, the average of the liquidity shocks becomes smaller and the variance increases. Since this is the case even when the platform usage cost is common to all dealers, when we let the period be a business week and not a day, the estimates suggest that the platform usage costs differ across dealers, but not the liquidity shocks. One explanation for this could be that investors attach different values to maintaining a close business relationship with different dealers.

In addition, we verify that our estimates are not driven by occasionally large trade sizes (see Appendix Table 10). We do so because our model abstracts from trade size, since most trades are small and similar in size. However, investors occasionally trade large amounts, and if they do, it is more likely that they trade bilaterally than on the platform (see Appendix Table 4). To test that our estimates are robust to these rare occasions, we re-estimate the model on a subsample of trades, excluding the 5% largest trades of an investor who trades more than one time. Investors who trade a

single time do not trade large amounts.

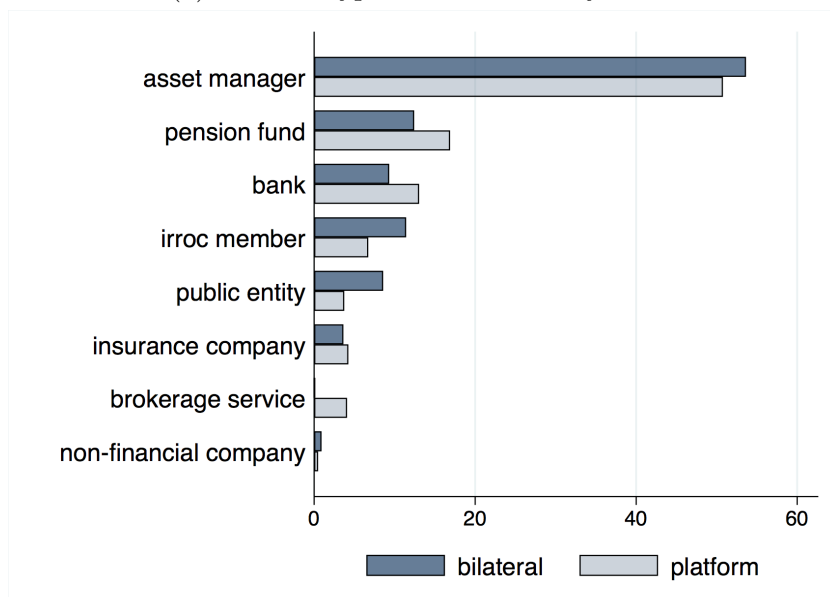
Finally, we test how sensitive our estimates are to the assumption that the dealer extracts all trade surplus in the bilateral negotiation and sets a yield that equals the investor's willingness to pay. For this, we rely on the extended model in Appendix B.2. We first re-estimate all model parameters imposing a positive bargaining power $\phi \in \{0.1, 0.25, 0.5\}$. For this, we extend the estimation procedure of the benchmark model, which is explained in Appendix C.2. The key difference is that we can no longer back out the dealer's value for the bond from the markup equation once we have estimated all other model parameters. Instead, the estimation procedure finds the implied value for the dealer for each constellation of parameters until it finds the set of parameters for which all moments are matched. Then, with the parameter estimates of the extended model, we repeat the counterfactual exercises to test the robustness of our main welfare findings.

As expected, we find that the parameter most sensitive to the choice of the investor's bargaining power ϕ is the investor's average willingness to pay for the bond, i.e., the average liquidity shock (see Appendix Table 11).²⁸ The effects on total welfare and the welfare decomposition remain essentially unchanged (see Appendix Figure A5).

²⁸We would expect that a buying investor who has some bargaining power would pay a price that is lower than her willingness to pay, and vice versa for the selling investor. This is true for the point estimates of all investor groups except for selling retail investors. For this investor group, standard errors are relatively large because retail investors do not sell as often in a day.

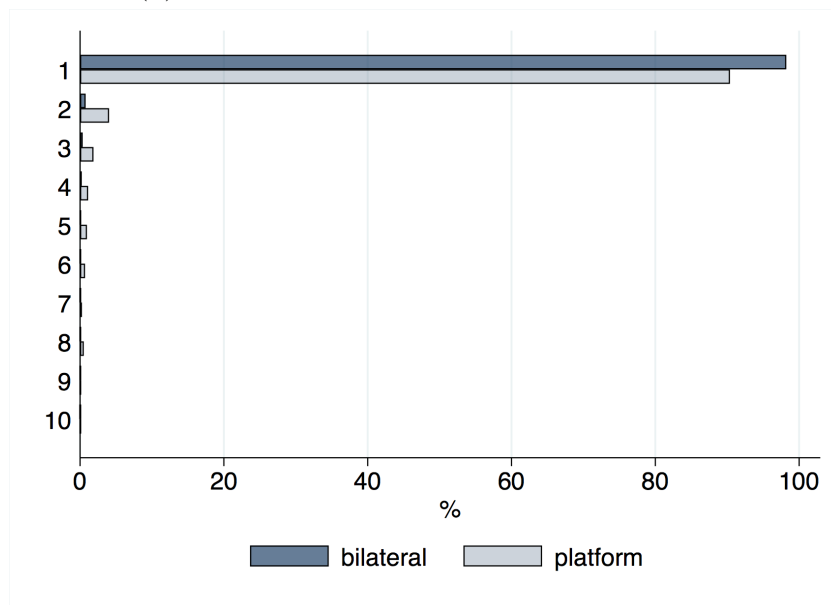
Appendix Figure A1

(a) Investor types and where they trade



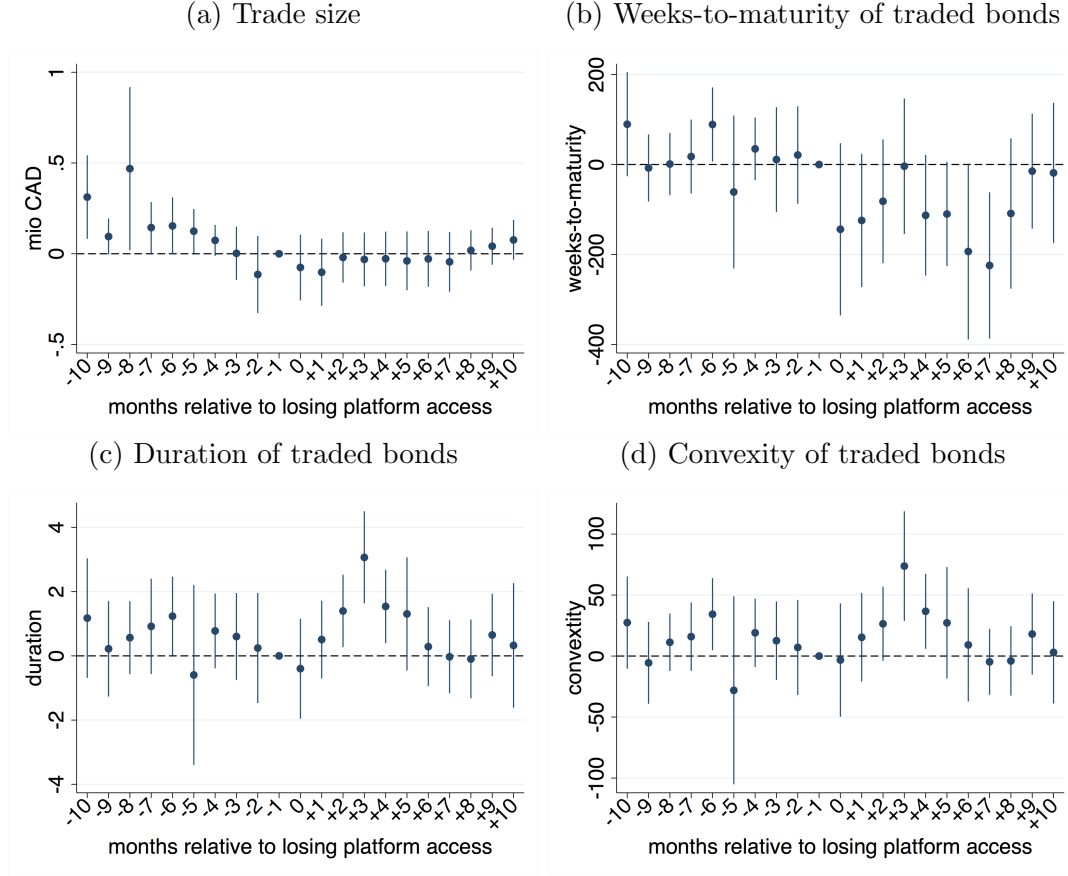
Appendix Figure A1a shows how much each investor group trades on the platform versus bilaterally as a percentage of the total amount investors trade.

(b) Number of dealers per investor



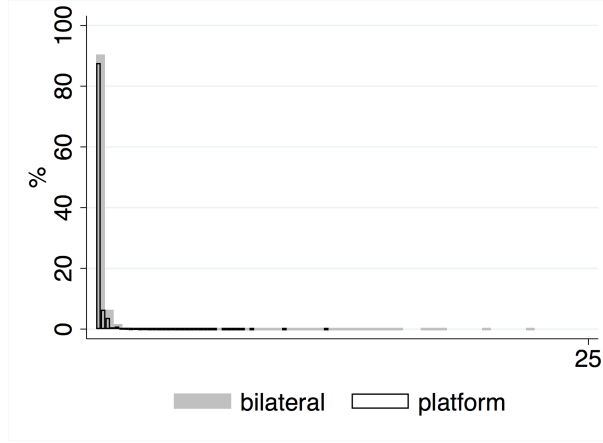
Appendix Figure A1b shows how many investor IDs trade with 1,...,10 primary dealers as a fraction of all IDs. Almost all trade with a single dealer. When restricting the sample to investors with LEIs, the fraction of IDs that trade with a single dealer decreases to about 70%. This is still high, considering that these investors are very active traders in this market.

Appendix Figure A2: Event study—observable trade behavior



These figures visualize changes in observable trade behavior when the investor loses platform access. They show the β_m estimates and the 95% confidence intervals of regressions that are similar to the event study regression (2), but with outcome variables that capture trade behavior. For Figure A2a we regress $quantity_{thsi} = \zeta_i + \sum_{m=M_i^-}^{M_i^+} \beta_m D_{mi} + \zeta_{th} + \zeta_s + \zeta_j + \epsilon_{thsi}$ to see whether the amount traded changes. Figures A2b to A2d illustrate whether the investor trades bonds with different characteristics; namely, the length to maturity, the duration (which approximates the bond's price sensitivity to changes in interest rates), and the convexity (which measures by how much the duration of the bond changes as interest rates change). In these regressions, we exclude the security fixed effect because it would absorb any characteristic of the bond. All standard errors are clustered at the investor level. The graphs look similar when we look at the number of dealers with whom investors trade in a month and how often or how much investors trade monthly.

Appendix Figure A3: Distribution of trade sizes



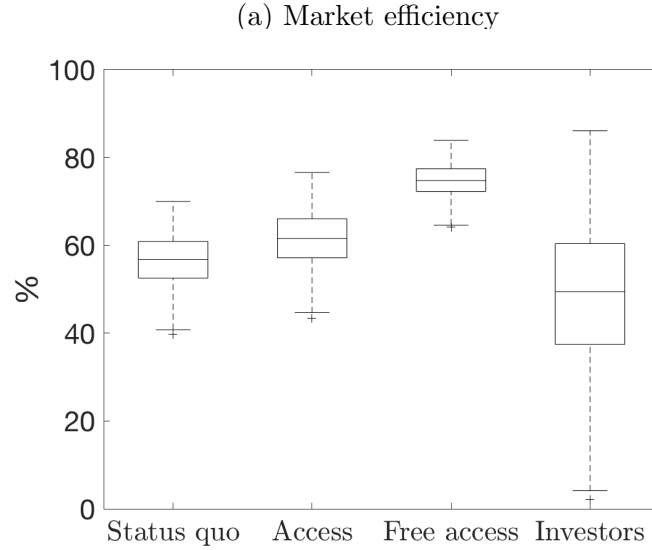
Appendix Figure A3 shows a probability density histogram of trade sizes in bilateral and platform trades. Trade size is measured in million C\$.

Appendix Figure A4: Number of times an investor trades in a day

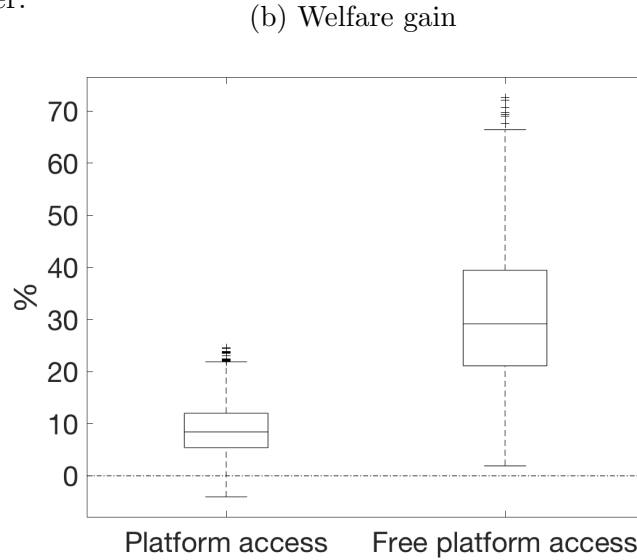


Appendix Figure A4 shows a probability density histogram of the number of times an investor trades—i.e., either sells or buys—in a day. The graph is similar when counting how many trading venues (bilateral vs. platform) an investor uses in a day. We see that in more than 95% of the cases, the investor only trades one time, and that on the rare occasions when the investor trades multiple times, she typically trades bilaterally and on the platform. Our model abstracts from these rare events.

Appendix Figure A5: Welfare analyses of the extended model ($\phi = 0.5$)



Appendix Figure A5a is the equivalent to Figure 8, but for the extended model with $\phi = 0.5$. It shows the distributions of daily welfare W_t , defined in (9), as the percentage of what could be achieved in the first best in four settings: the status quo, the counterfactuals in which all investors have platform access at the estimated platform usage costs and for free, and the counterfactual in which investors trade with one another.



Appendix Figure A5b is the equivalent to Figure 7, but for the extended model with $\phi = 0.5$. It illustrates by how much welfare increases when making platform access universal or free. In both cases, it shows the distribution of the percentage change in welfare, $\Delta W_t / W_t * 100\%$, over days with W_t and ΔW_t , as in (9) and (10).

Appendix Table 1: Sample restrictions

Restrictions	Sample size	Size ↓ in %
All dealer-to-investor trades	1,948,764	
w/o extreme yields	1,914,031	1.78%
w/o in-house trading	1,668,520	12.82%
w/o errors in trading venue	1,620,148	2.89%
w/o out of business hours	1,523,037	5.99%
w/o false investor-type indicator	1,517,714	0.34%
w/o trades after July 2019 (model only)	1,346,462	11.28%
w/o non primary dealers (model only)	1,252,718	6.96%
w/o one of the primary dealers (model only)	1,139,412	9.04%
w/o trades prior announcement (model only)	1,003,542	11.92%

Appendix Table 1 summarizes how we restrict the raw data. We exclude extreme yields and trades by institutions that are likely reported as institutional investors but are retail or vice versa. Further, we exclude in-house trades, trades that are not realized on CanDeal or bilaterally, and trades that occur out of business hours. For the structural estimation, we focus on trades with primary dealers only. We exclude trades after July 2019 because our auction data do not cover the second half of 2019. Lastly, we exclude trades on auction dates prior to the auction announcement and drop one primary dealer due to data reporting.

Appendix Table 2: Yields are better on the platform

	(1)		(2)	
platform	0.282	(0.0331)	0.0795	(0.0310)
constant	-0.296	(0.00934)	-0.281	(0.00690)
investor fixed effect (ζ_i)	—		✓	
Observations	1,193,999		806,473	
Adjusted R^2	0.169		0.523	

Appendix Table 2 shows that yields on the platform are better than those off the platform. For this, we regress the markup $(y_{thsi j} - \theta_{ths})^+$, as defined in (1), on an indicator variable that assumes value 1 if the trade realizes on the platform ($platform_{thsi j}$), hour-day (ζ_{th}), security (ζ_s), and dealer (ζ_j) fixed effects. In column (2) we add investor (ζ_i) fixed effects. Standard errors are in parentheses and clustered at the investor level in column (2)

Appendix Table 3: Effect of trade size on yields

buy	1.313	(0.135)
θ	0.988	(0.000690)
quantity	0.0246	(0.0388)
quantity ²	-0.0177	(0.0122)
quantity ³	0.00120	(0.000737)
constant	0.881	(0.102)
Observations	806,564	
Adjusted R^2	0.998	

Appendix Table 3 shows the estimation results when regressing the trade yield ($yield_{thsi j}$) on an indicator variable that shows the size of the trade ($buy_{thsi j}$), the market value (θ_{ths}), and a function of trade size, $\sum_{p=1}^3 \delta_p (quantity_{thsi j})^p$, in addition to hour-day (ζ_{th}), security (ζ_s), dealer (ζ_j), and investor (ζ_i) fixed effects. The findings suggest that trade size is not driving the yield, since all of the coefficient multiplying quantity are statistically insignificant. Standard errors are in parentheses and clustered at the investor level.

Appendix Table 4: Trade size and venue choice

	(1)	(2)	(3)
quantity	-0.101 (0.00104)	-0.0373 (0.00911)	-0.0213 (0.0163)
constant	0.353 (0.000452)	0.330 (0.00157)	0.331 (0.00207)
investor fixed effect (ζ_i)	—	✓	✓
Observations	1,234,945	784,809	738,344
Adjusted R^2	0.008	0.462	0.468

Appendix Table 4 shows whether institutional investors trade different amounts on versus off the platform. Column (1) gives the estimation results when regressing an indicator for whether trade realizes on or off the platform on the trade size (quantity). In column (2), we add an investor fixed effect (ζ_i). In column (3), we exclude the 5% largest trades of an investor to show that the statistically significant negative correlation between platform and quantity is driven by occasional large trades. Our interpretation is that in these rare cases, investors prefer to trade bilaterally with their dealer. This could be because their dealer works as an insurance for rainy days and offers a better deal than other dealers with whom the investor does not maintain a close business relationship. In this paper, we focus on regular small trades. Standard errors are in parentheses and clustered at the investor level in columns (2) and (3).

Appendix Table 5: Bilateral yields and the supply shocks

	(1) buys		(1) sells	
θ_{ths}	0.668	(0.00249)	0.749	(0.00195)
$won_{\tilde{t}j}$	-0.0000199	(0.0000588)	-0.0000567	(0.0000458)
constant	47.73	(0.369)	37.82	(0.288)
Observations	192,342		192,004	
Adjusted R^2	0.724		0.681	

Appendix Table 6: Quotes and the supply shocks

	(1) buys		(1) sells	
θ_{ths}	0.0224	(0.000990)	0.0351	(0.000963)
$won_{\tilde{t}j}$	0.000489	(0.0000234)	0.000291	(0.0000226)
constant	144.4	(0.147)	142.7	(0.142)
Observations	192,342		192,004	
Adjusted R^2	0.272		0.237	

Here we provide evidence that dealers do not adjust bilateral yields when hit by an unexpected supply shock, defined in (8). For Appendix Table 5, we regress the bilateral yield of a trade between investor i and dealer j in hour h of date t with security s on the bond's market value, and the dealers' supply shocks, as well as dealer, investor, and date fixed effects— $y_{thsij} = \alpha + \beta\theta_{ths} + \gamma won_{\tilde{t}j} + \zeta_j + \zeta_i + \zeta_t + \epsilon_{thsij}$ —for buying investors in column (1) and selling investors in column (2). Appendix Table 6 shows the analogous results when replacing the bilateral yield in both regressions with the platform quote q_{tj} . All yields are in basis points and $won_{\tilde{t}j}$ is in million C\$. Standard errors are in parentheses.

We find that both $won_{\tilde{t}j}$ coefficients in Appendix Table 5 are close to 0 and insignificant. The estimates imply that when $won_{\tilde{t}j}$ increases by one standard deviation, the change in the yield lies in (-0.0145 bps, +0.0102 bps) for buying investors and in (-0.0159 bps, +0.0036 bps) for selling investors. When clustering standard errors, either at the dealer or investor level or both, these intervals become even tighter. In contrast, the dealer adjusts her quotes according to Appendix Table 6. Both findings are in line with our model assumptions.

Appendix Table 7: Robustness of $1/\sigma$ w.r.t. the instrument

Specification	OLS	IV1	IV2	OLS	IV1	IV2
quote coefficient ($1/\sigma$)	0.014 (0.005)	1.467 (0.259)	1.287 (0.246)	-0.093 (0.011)	1.612 (0.175)	2.050 (0.236)
dealer fixed effect (ζ_j)	✓	✓	✓	—	—	—
Observations	8,492	8,492	8,492	8,492	8,492	8,492
Adjusted R^2	0.804	0.805	0.805	0.040	0.040	0.040

Appendix Table 7 shows how the σ parameter changes depending on the instrument we use. Specifically, it gives the point estimate of regression (16) in Appendix C.2, which is the inverse of σ . The second and fifth columns show the OLS estimates, first including a dealer fixed effect and then excluding it. In the third and sixth columns, we instrument the relative quote with the amount the dealer won on the most recent auction day. The fourth and seventh columns show the estimate using the unexpected supply shocks as instruments (as reported in the text). Standard errors are in parentheses.

Appendix Table 8: Estimates (median across days) when $y_{ti}^G = \theta_t + \nu_{ti}^G$

Buys	$\hat{\mu}^I$	$\hat{\mu}^R$	$\hat{\sigma}^I$	$\hat{\sigma}^R$	\hat{c}	$1/\hat{\sigma}$	$\hat{\eta}$
	-0.87 (0.13)	-2.97 (0.73)	2.56 (0.12)	5.05 (0.94)	3.34 (0.18)	1.29 (0.94)	+174.13
Sells	$\hat{\mu}^{I*}$	$\hat{\mu}^{S*}$	$\hat{\sigma}^{I*}$	$\hat{\sigma}^{S*}$	\hat{c}_t^*	$1/\hat{\sigma}$	$\hat{\eta}^*$
	+0.97 (0.14)	+2.04 (0.64)	2.62 (0.11)	4.50 (0.98)	3.46 (0.18)	1.29 (0.94)	-174.23

Appendix Table 8 is the analogue to Table 1, but here we assume that $y_{ti}^G = \theta_t + \nu_{ti}^G$. It shows the median over all days of all point estimates per investor group G , in addition to the implied elasticity of demand ($\hat{\eta}$) and of supply ($\hat{\eta}^*$) on the platform, averaged across days and dealers. The corresponding medians of the standard errors are in parentheses. All estimates are in bps.

Appendix Table 9: Estimates when the platform usage cost is dealer specific

Buys	$\hat{\mu}^I$	$\hat{\sigma}^I$	\hat{c}_1	\hat{c}_2	\hat{c}_3	\hat{c}_4	\hat{c}_5	\hat{c}_6	\hat{c}_7	\hat{c}_8	\hat{c}_9	$\hat{\eta}$
	-0.77	2.67	3.04	3.49	2.71	4.06	3.74	2.95	1.67	3.23	3.56	+173.82
Sells	$\hat{\mu}^{I*}$	$\hat{\sigma}^{I*}$	\hat{c}_1^*	\hat{c}_2^*	\hat{c}_3^*	\hat{c}_4^*	\hat{c}_5^*	\hat{c}_6^*	\hat{c}_7^*	\hat{c}_8^*	\hat{c}_9^*	$\hat{\eta}^*$
	+0.89	2.76	3.42	3.72	2.82	4.00	3.66	3.34	1.97	1.97	3.42	-173.95

Appendix Table 9 is similar to Table 1, but here we assume that the platform usage cost is dealer specific and count a business week rather than a day as a period t . The table shows the median over all weeks of all point estimates for institutional investors I , in addition to the implied elasticity of demand ($\hat{\eta}$) and of supply ($\hat{\eta}^*$) on the platform, averaged across days and dealers. All estimates are in bps.

Appendix Table 10: Estimates (median across days)

Buys	$\hat{\mu}^I$	$\hat{\mu}^R$	$\hat{\sigma}^I$	$\hat{\sigma}^R$	\hat{c}	$1/\hat{\sigma}$	$\hat{\eta}$
	-0.79 (0.13)	-2.95 (0.79)	2.79 (0.10)	5.14 (0.92)	-3.34 (0.16)	1.36 (0.25)	+184.51
Sells	$\hat{\mu}^{I*}$	$\hat{\mu}^{S*}$	$\hat{\sigma}^{I*}$	$\hat{\sigma}^{S*}$	\hat{c}_t^*	$1/\hat{\sigma}$	$\hat{\eta}^*$
	+0.91 (0.13)	+1.99 (0.67)	2.86 (0.10)	4.48 (0.94)	-3.46 (0.17)	1.36 (0.25)	-184.51

Appendix Table 10 shows the estimation results when restricting the sample to trades of regular trade sizes, excluding the 5% largest trades of investors who trade more than once. Analogous to Table 1, it shows the median over all days of all point estimates per investor group G , in addition to the implied elasticity of demand ($\hat{\eta}$) and of supply ($\hat{\eta}^*$) on the platform, averaged across days and dealers. All estimates are in bps.

Appendix Table 11: Estimates of the extended model ($\phi = 0.1$)

Buys	$\hat{\mu}^I$	$\hat{\mu}^R$	$\hat{\sigma}^I$	$\hat{\sigma}^R$	\hat{c}	$1/\hat{\sigma}$	$\hat{\phi}$
	-0.96	-3.31	3.08	5.61	3.38	1.29	+174.57
Sells	$\hat{\mu}^{I*}$	$\hat{\mu}^{S*}$	$\hat{\sigma}^{I*}$	$\hat{\sigma}^{S*}$	\hat{c}_t^*	$1/\hat{\sigma}$	$\hat{\phi}^*$
	+0.44	+2.24	2.37	5.00	3.69	1.29	-179.59

Appendix Table 12: Estimates of the extended model ($\phi = 0.25$)

Buys	$\hat{\mu}^I$	$\hat{\mu}^R$	$\hat{\sigma}^I$	$\hat{\sigma}^R$	\hat{c}	$1/\hat{\sigma}$	$\hat{\phi}$
	-1.23	-4.17	3.59	6.72	3.27	1.29	+174.42
Sells	$\hat{\mu}^{I*}$	$\hat{\mu}^{S*}$	$\hat{\sigma}^{I*}$	$\hat{\sigma}^{S*}$	\hat{c}_t^*	$1/\hat{\sigma}$	$\hat{\phi}^*$
	+0.71	+2.90	2.79	5.98	3.69	1.29	-178.67

Appendix Table 13: Estimates of the extended model ($\phi = 0.5$)

Buys	$\hat{\mu}^I$	$\hat{\mu}^R$	$\hat{\sigma}^I$	$\hat{\sigma}^R$	\hat{c}	$1/\hat{\sigma}$	$\hat{\phi}$
	-2.01	-6.59	5.13	10.07	3.06	1.29	+174.44
Sells	$\hat{\mu}^{I*}$	$\hat{\mu}^{S*}$	$\hat{\sigma}^{I*}$	$\hat{\sigma}^{S*}$	\hat{c}_t^*	$1/\hat{\sigma}$	$\hat{\phi}^*$
	+1.53	+4.81	4.07	8.94	3.62	1.29	-177.11

Appendix Tables 11–13 are similar to Table 1, but of the extended model in which the investor has a bargaining power of $\phi = 0.1$, $\phi = 0.25$, or $\phi = 0.5$. The tables show the median over all days of all point estimates per investor group G , in addition to the implied elasticity of demand ($\hat{\eta}$) and of supply ($\hat{\eta}^*$) on the platform, averaged across days and dealers. All estimates are in bps.