DEALER INTERMEDIATION BETWEEN MARKETS

Peter G. Dunne
Central Bank of Ireland

Harald Hau
University of Geneva and Swiss Finance Institute

Michael J. Moore
Warwick Business School

Abstract
We develop a dynamic model of dealer intermediation between a monopolistic customer–dealer market and a competitive interdealer limit order market. Dealers face inventory constraints and adverse selection. We characterize the optimal quote setting and inventory management behavior for both markets in closed form and reveal how price setting in one market segment influences quote behavior in the other. The framework is used to explore market stability issues of the two-tier market structure and delivers testable predictions about how the dispersion of retail prices is related to the state of the interdealer limit order book. Data from the European sovereign bond market is used to test for inventory related retail price dispersion. (JEL: G24, G14)

1. Introduction

Dealers are intermediaries between different market segments. A dealer maintains a network of customer relationships and simultaneously participates in an interdealer market which allows her to manage her inventory. In the customer segment, the dealer typically has some market power because her clients face search costs and do not have direct access to the wholesale or interdealer market. Interdealer markets on the other hand are often highly competitive and have become dominated by electronic limit order books.

The editor in charge of this paper was George-Marios Angeletos.

Acknowledgments: We thank Euro MTS for their generous access to the data. We also thank KX-Systems, Palo Alto, and their European partner, First Derivatives, for providing their database software Kdb. We are also grateful for comments from participants at the EFA 2010 meetings and various other seminars. We thank Andrey Zholos and James Waterworth for excellent research assistance. A special thanks goes to Thierry Foucault, Denis Gromb, Christine Parlour, and Albert Menkveld for their comments. Moore thanks the UK Economic and Social Research Council for financial assistance under grant number ES/H041524/1.

E-mail: peter.dunne@centralbank.ie (Dunne); prof@haraldhau.com (Hau); michael.moore@wbs.ac.uk (Moore)

Journal of the European Economic Association October 2015 13(5):770–804

DOI: 10.1111/jeea.12118

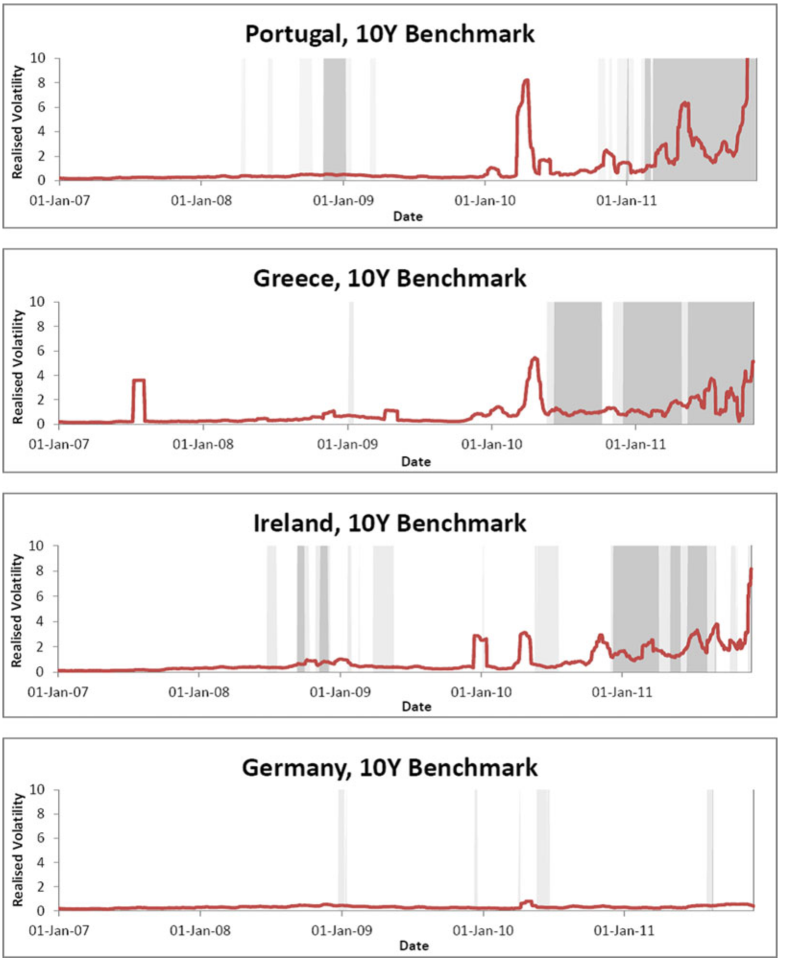

FIGURE 1. A 20-day moving average of realized intra-day volatility based on return measurement at 15 minute intervals for various 10-year benchmark bonds. The dark (light) gray background shading highlights (partial) interdealer market breakdown for periods in which the daily trading volume falls below the 3% (resp. 10%) average trading volume from January 2007 to July 2007 prior to the European sovereign bond crisis.

Such two-tier market structures have proven very fragile in the recent financial crisis as exemplified by the experience of the European sovereign bond market—the world's largest fixed income market. In Figure 1, the dark (light) shaded periods show a liquidity shortfall of at least 97% (resp. 90%) percent in the interdealer trading platform during the recent European sovereign debt crisis. The market breakdown for Portuguese, Greek, and Irish 10-year benchmark bonds coincides (unlike for German government bonds) with a considerable increase in realized volatility for the respective bond returns. What explains such fragility of a two-tier market structure of dealer–client relationships and interdealer trading? How can such a two-tier market structure be adequately represented in a dynamic framework?

This paper develops a dynamic model of dealer intermediation between a competitive interdealer limit order market and the dealer–customer segment. We fully characterize the dynamic limit order equilibrium in the interdealer market and derive the optimal retail quotes in the dealer–customer segment in closed form. Compared to previous work on limit order markets, our framework puts more structure on the dealer problems by modeling two market segments. Yet, we obtain a tractable equilibrium solution, which provides new insights into the stability of dealership market structures—in particular about their resilience in times of high market volatility.

For the most part, limit order markets like the interdealer market have been studied in isolation.[1] Following Glosten (1994), the ask (bid) side price schedule of a competitive limit order market has typically been characterized in terms of the so-called upper (lower) tail expectation. Heterogeneity of private asset valuations will tend to lower spreads and flatten the limit order supply schedules, whereas adverse selection risk has the opposite effect. As limit order markets provide dealers with a large menu of trading strategies, it has proven difficult to obtain simple closed form solutions in a fully dynamic setting. As a consequence, important policy issues related to market stability remain largely outside the realm of microeconomic analysis.

The tractability in our framework is achieved by using a very economic structure to represent the dealership problem. We assume that dealers face inventory constraints, which condition their choice between limit and market orders. Importantly, all trading benefits in the interdealer segment are restricted to inventory rebalancing among (ex-ante) identical dealers. Adverse selection risk follows from changes in the aggregate customer demand rather than from private information about market fundamentals; adverse selection risk is thus tied to the volatility of the asset fundamentals. Our model is therefore particularly pertinent for the sovereign bond market in which asymmetric private information should be less relevant.

We provide new insights into the fragility of the two-tier market structure, the interdependence between the interdealer and the retail segments, and the effect of regulatory measures like security transaction taxes on market stability. First, dealer intermediation in a two-tier market structure renders the existence of a market equilibrium precarious because of likely market breakdown in the interdealer segment. The inside spread in the wholesale market is shown to be the difference between the expected loss from adverse selection and the dealer's benefit of reaching a balanced inventory state. As the latter benefits are limited because of (ex-ante) dealer homogeneity, adverse selection risk may easily outweigh the rebalancing benefits and compromise equilibrium existence. We are able to characterize the point of market breakdown in the two-tier market. In a one-tier market dealers can also hope to trade against a large group of clients which in the two-tier-structure are "captured" by other dealers. The dispersion of private valuations of these clients (available for trading in a consolidated one-tier market) contribute to much bigger gains from trade (compared to pure interdealer trades) and prevents market breakdown even under increased levels of adverse selection.

1. See Parlour and Seppi (2008) for a recent survey.

Second, the fragility of the two-tier market structure is further increased by the interdependence between the interdealer and retail market segment. While outside the scope of most microstructure models, our framework captures a key feature of dealer market intermediation: the optimal retail quotes typically depend on the dealer's rebalancing costs in the interdealer segment. Higher interdealer spreads increase retail spreads for certain inventory states, which in turn increases the adverse selection component of retail order flow, which is passed on to the interdealer market for rebalancing. More adverse selection risk then increases interdealer spreads further and feeds back into higher retail spreads—thus creating a "feedback loop", which represents a further aspect of fragility of the two-tier market structure.

Third, any trade execution cost or security transaction tax in the interdealer market will further decrease the volatility threshold of market breakdown through two channels. First, they reduce the already limited trading benefits between dealers. Second, higher rebalancing costs increase optimal quote spreads in the dealer–customer segment and—through the feedback loop—increase the adverse selection risk of liquidity provision in the interdealer market. Our framework provides important insights as to how transaction costs affect the robustness of the two-tier market structure through a multiplier effect.

A second policy concern for dealer markets is the dispersion of retail prices (Harris and Piwowar 2006; Green, Hollifield, and Schurhoff 2007). Our framework delivers specific predictions about the retail price distribution and how it depends on the state of the interdealer market. Retail price dispersion may reflect benign inventory management concerns of constrained dealers or alternatively their price discrimination across different customer types. Our dealer market model helps to identify when price dispersion is driven by the former rather than the latter. To illustrate this aspect, we provide an application to the European sovereign bond market. Using synchronized transaction data from the interdealer and retail segment, we show that the average retail market quality on the ask (bid) side increases whenever the interdealer limit order book becomes deep for the best bid (ask) limit order. Inventory-motivated retail price dispersion can be identified directly from the state of the interdealer limit order book, because the latter reflects the inventory dispersion across all dealers. By contrast, customer-type-based retail price discrimination should be invariant to the state of the interdealer limit order book—a hypothesis rejected for the European sovereign bond market.

A shortcoming of our baseline model is the monopolistic market structure in the retail segment, which ties each customer to a single dealer and prevents the former from shopping for better quotes. We therefore extend the framework by assuming that only a share of customers is captured by a dealer, whereas "sophisticated customers" shop for the best deal. In particular, sophisticated customers undertake transactions only with those dealers which—due to inventory imbalances—feature the most favorable reservation prices, and we assume that such customers extract all the transaction rents. We interpret a higher share of sophisticated clients as a more competitive retail market structure and explore its effects on market quality. More customer sophistication has the consequence that extreme inventory states and their favorable quotes attract more customers and thus reduce the trading benefits in the interdealer market. This renders the two-tier market structure even more fragile and market breakdown occurs at a lower level of volatility. Competition in the retail segment and market stability are therefore inversely related.

The following section describes the contribution of the literature before we present the baseline model and its solution in Section 3. Section 4 explores the welfare benefits of the two-tier market relative to a pure retail market in which no central interdealer market exists. Section 5 extends the analysis to allow for dealer competition. The empirical Section 6 tests a specific prediction of our cross-market intermediation model; namely that the bid-side (ask-side) market depth in the interdealer market determines the average ask-side (bid-side) retail quote quality. Section 7 discusses limitations of the analysis and Section 8 concludes.

2. Analytical Tractability and the Literature

Our work is related to research on limit order markets recently surveyed by Parlour and Seppi (2008). The microstructure literature usually considers the limit order market in isolation based on exogenous trader arrival (Goettler, Parlour, and Rajan 2009; Rosu 2014). Yet even in this stylized setting, analytical solutions are typically not available because traders face so many endogenous choices about the order type (limit versus market order), and limit price and quantity as a function of the entire order book. Generally, trader heterogeneity or asymmetric information make the equilibrium analysis intractable.

We take a different approach by deriving the dealers' trading needs in the limit order market directly from a retail market process and—surprisingly—this aspect contributes to deeper economic structure as well as to more tractability of the dynamic limit order market equilibrium. Only the customer arrival process in the retail segment of the market is stochastic, but participation in the interdealer market by all dealers is continuous. We highlight three key assumptions which allow for an analytical solution.

First, as in Foucault (1999), we assume that a common value process x_t evolves on a binomial tree with private customer values for the asset distributed (uniformly) around this value. The adverse selection risk for the liquidity supplier consists in limit order provision without knowledge of the next innovation $\Delta x_{t+1} = x_{t+1} - x_t \in \{-\epsilon, +\epsilon\}$ in the market demand. This representation is particularly parsimonious and ties the adverse selection risk simply to the volatility of the common value process. High price volatility becomes a sign of market stress and should be correlated with market breakdown—as observed in various crisis episodes.[2]

Second, the interdealer market consists only of (ex-ante) identical dealers facing a simple inventory constraint. Inventory constraints condition both the benefits of trade among dealers as well as the optimal order submission choice. Hitting an inventory

2. See Ranaldo (2004) for evidence that the inside limit order bid–ask spread indeed increases in price volatility.

limit always requires a dealer to rebalance via a market order—whereas dealers choose limit orders otherwise. Compared to previous models with endogenous order type choice (Kumar and Seppi 1994; Foucault, Kadan, and Kandel 2005; Kaniel and Liu 2006; Goettler, Parlour, and Rajan 2009), the optimal order submission strategy is thus relatively easy to characterize. In particular, the private benefits of trade are restricted to rebalancing and tied to the concavity of the dealer's value function across different inventory states. To simplify the analysis further, we also assume that all transactions occur in one unit of the traded asset and eligible inventory states are restricted to three states: $s = -1, 0, +1$. In this parsimonious structure, the limit order market equilibrium can be characterized in terms of only two endogenous variables—namely the rebalancing benefit ∇ (of moving to a balanced inventory state) and the (inside) interdealer market spread S. Both variables depend on the adverse selection risk embodied in the customer order flow coming from the retail segment.

Third, the dealers' trading needs in the interdealer segment are endogenous and come from customer order flow in the retail segment. We derive this customer order flow directly for the dealers' profit maximizing retail quote behavior. The private asset value distribution of clients around the common value x_t consists of a uniform distribution; thus we obtain a linear closed-form solution for the optimal retail quotes as a function of a dealer's inventory state. The customer arrival process is stochastic and customers seek dealer quotes only from "their dealer" and (in our baseline model) they do not shop around for alternative dealer quotes. It is easy to show that the dealers' optimal dynamic retail quote behavior features so-called "inventory shading": they lower customer quotes on the bid (ask) side in the case of positive (negative) inventory imbalances. Importantly, the degree of inventory shading depends again on the interdealer spread S and the value concavity parameter ∇. Both the retail market and the interdealer market segments are therefore interdependent; dealer intermediation between the two markets is predicated on a joint equilibrium linking both markets.

An extended literature has explored the role of dealers as arbitrageurs between different markets. Unlike the dealer intermediation for the same asset between a wholesale and a retail market, such cross-asset arbitrage should be more sensitive to funding capital and its withdrawal in a financial crisis (Gromb and Vayanos 2002, 2010; Brunnermeier and Petersen 2009; Lagos, Rocheteau, and Weill 2009; Rinne and Suominen 2009; Duffie and Strulovici 2012; He and Krishnamurthy 2012). By contrast, arbitrage between different market participants for the same asset does not generate large net funding needs and market breakdown is likely to result from other forces than a withdrawal of funding liquidity.[3]

Other work has modeled OTC structures as a result of trading restrictions or market inattention in which dealers can trade at any moment whereas customers trade infrequently (Duffie 2010). While such assumptions allow an easy comparison with a fully integrated one-tier market (by lifting the trading restrictions), they abstract from a key economic aspect specific to the two-tier dealership structure: dealers can quote

3. Total funding liquidity for our market model is zero as dealers are as likely to hold liquidity-generating short positions as liquidity-consuming long positions.

different prices in different market segments as customers are typically excluded from interdealer segment. We consider these aspects crucial to the two-tier market structure and make it the focus of our modeling approach.

3. A Model of Cross-Market Intermediation

Most financial markets feature a dual (or two-tier) market structure in which dealers maintain a network of client relationships (B2C) and have access to an interdealer (B2B) trading platform. Clients are excluded from participation in the B2B market and have to transact directly with a dealer. Dealer intermediation thus occurs across market segments of different competitiveness. The interdealer market is typically highly competitive and organized as a limit order market, whereas client relationships and client search costs might provide the dealer with some market power in the dealer–customer segment.

3.1. Assumptions

Dealers face a stochastic arrival process for potential customers with uncertain private values. The customer arrival process has the following structure.

ASSUMPTION 1 (*Customer Arrival and their Reservation Prices*). Each period a dealer faces customer requests for buy (sell) quotes with a constant probability q. Let R^a and R^b denote the private customer values such that the customer buys if $R^a > \hat{a}$ and sells if $R^b < \hat{b}$, where the requested ask and bid prices $(\hat{a}, \hat{b})$ are set one period ahead. Private customer values have a uniform distribution with density d over the interval $[x_{t+1}, x_{t+1} + d^{-1}]$ and $[x_{t+1} - d^{-1}, x_{t+1}]$ for the ask and the bid, respectively. The midprice x_{t+1} is a stochastic martingale process known to all dealers only at time $t+1$. For simplicity we choose $\Delta x_{t+1} = x_{t+1} - x_t \in \{-\epsilon, +\epsilon\}$ with corresponding probabilities $(0.5, 0.5)$ and assume an upper bound for volatility with $\epsilon < \bar{\epsilon}$. All transactions concern a quantity of one unit.

Assumption 1 characterizes the competitive situation of each dealer in the B2C market segment. More unfavorable client quotes reduce (linearly) the chance of customer acceptance. The customer arrival probability q is exogenous, identical for the bid and ask side, and does not depend on a dealer's quote quality. The martingale process x_t represents the common value component of the asset from which the private valuations of bid- and ask-side clients symmetrically deviate. The private value assumption implicitly grants dealers a certain degree of monopolistic market power that depends on the parameter d. A smaller d increases the monopolistic rents a dealer can earn from the dealer–client relationship. The exogenous distribution of customer reservation prices excludes any strategic interaction between dealers, whereby the pricing behavior of a single dealer alters the customer demand for another dealer. Each dealer is assumed to be atomistic. We also assume that the parameter d is constant over time and does not depend on the volatility of the midprice process.[4]

It is assumed that dealers quote optimal ask and bid prices for period $t+1$ based on knowledge of the midprice x_t, but not yet based on the new realization x_{t+1}. Hence dealer-quoted customer prices incorporate demand shocks only with a one-period delay. This subjects dealers to an adverse selection problem that widens spreads. The adverse selection risk increases in the variance ϵ^2 of the midprice process x_t. For simplicity we require that the shift to the reservation price distribution is bounded by $\bar{\epsilon}$ so that the ex-ante optimal B2C quotes in all inventory states are still on the support of this distribution at time $t+1$.[5]

It is useful to denote standardized ask and bid quotes by $a = \hat{a} - x_t$ and $b = \hat{b} - x_t$, respectively.[6] Standardized quotes represent the quoted dealer prices relative to the current expected midprice $x_t = \mathcal{E}(x_{t+1})$. We also define cumulative density functions for the acceptance of a dealer quote as

$$\begin{aligned}
F^a\left(R^a \geq \hat{a}\right) &= F^a\left(R^a - x_{t+1} \geq \hat{a} - x_{t+1} = a - \Delta x_{t+1}\right) \\
&= 1 - ad + d\Delta x_{t+1} \\
F^b(R^b \leq \hat{b}) &= F^b(R^b - x_{t+1} \leq \hat{b} - x_{t+1} = b - \Delta x_{t+1}) \\
&= 1 + bd - d\Delta x_{t+1},
\end{aligned}$$

respectively. A higher dealer ask price a, for example, reduces the quote acceptance linearly. The term $d\Delta x_{t+1}$ captures changes in the acceptance probability resulting from the exogenous evolution of the reservation price distribution.

For the purpose of inventory management, dealers can resort to an interdealer market with a spread $S = \hat{A} - \hat{B} > 0$.

ASSUMPTION 2 (*Competitive Interdealer (B2B) Market*). Dealers have access to a fully competitive interdealer market and can (via market orders) buy inventory at the (best) ask price $\hat{A}$ and sell at the (best) bid price $\hat{B}$. The interdealer prices are cointegrated with the price process x_t with $\hat{A} = x_t + 0.5S$ and $\hat{B} = x_t - 0.5S$. We refer to standardized interdealer prices as $A = \hat{A} - x_t = 0.5S$ and $B = \hat{B} - x_t = -0.5S$, respectively and assume $0.5S \in [0, d^{-1} - 2\epsilon]$. The ask and bid (limit order) prices A and B are set competitively (i.e., equal a dealer's reservation price) by a large number of dealers distributed across all inventory levels. Interdealer transactions require order-processing costs of τ per transaction for liquidity providers.

4. In principle, the parameter d could also differ on the ask and the bid side of the market. This would give rise to asymmetric market power on the ask and bid side and allow for a richer asymmetric distribution of B2C quote behavior. For simplicity, we focus on the symmetric case.

5. A necessary and sufficient condition is that $\epsilon < 0.5d^{-1} - 0.25S$ where S denotes the B2B spread. The endogenously determined $S(\epsilon)$ is an increasing function in ϵ. Hence, the upper bound for volatility is implicitly defined as $\bar{\epsilon} = 0.5d^{-1} - 0.25S(\bar{\epsilon})$.

6. Hereafter, the expression "standardized quotes" means the deviation of the quote from the prevailing B2B midprice.

The interdealer market allows dealers to manage their inventory and adhere to their inventory constraints. The interdealer spread reflects all public dealer information about the price x_t. Order-processing costs are captured by the parameter τ.[7] We later explore the effects of security transaction taxes (STT) by permitting a change in τ.

ASSUMPTION 3 (*Dealer Objectives and Inventory Constraints*). A dealer chooses optimal B2C quotes $(\hat{a}, \hat{b})$ at the ask and bid side, respectively, in order to maximize the expected payoff under an inventory constraint that limits her inventory level to the three values $I = 1, 0, -1$. She is required to liquidate any inventory above 1 or below -1 immediately in the interdealer market. Let $0 < \beta = (1 + r)^{-1} < 1$ denote the dealer's discount factor for an interest rate on capital r. Let $n(I)$ be the number of dealers at each inventory level. We assume furthermore that the probability q of customer arrival in the B2C market is sufficiently small so that $0.5q < n(1)/n(-1) < 2/q$ holds.

In order to limit the number of state variables we allow for only three inventory levels. This choice greatly facilitates the exposition. Inventory constraints embody the idea that dealers work within managerially pre-set position limits during the course of trading.[8] Direct empirical evidence about the role of inventory constraints in dealer markets mostly relates to equity markets (Hansch, Naik, and Viswanathan 1998; Reiss and Werner 1998).

The condition on the arrival probability is needed to ensure that dealer rebalancing at the best B2B spread is always feasible to avoid a one-sided illiquidity problem in the B2B market. For a continuum of dealers, $n(I)$ can be interpreted as the probability mass of dealers in each inventory state. For a discrete dealer set, a highly nonsymmetric dealer distribution over the inventory states (with $n(1) = 0$ or $n(-1) = 0$) remains a small, but nonzero probability, which is neglected in the consecutive analysis. The sequence of trading is summarized in Figure 2.

3.2. A Dealer's Value Function

We denote a dealer's value function for the present value of all future expected payoffs by $V(s, x_t)$. The state variable $s = 1, 0, -1$ represents one of the three possible inventory values. Furthermore, let $p_{s_t s_{t+1}}$ denote the transition probability of state s_t in period t to state s_{t+1} in period $t + 1$. For three states, a total of nine transition probabilities characterize the transition matrix

$$\mathbf{M} = \begin{bmatrix} p_{12} + p_{11} & p_{10} & 0 \\ p_{01} & p_{00} & p_{0-1} \\ 0 & p_{-10} & p_{-1-1} + p_{-1-2} \end{bmatrix}.$$

7. Order-processing costs may be relatively important in interdealer markets as argued by Huang and Stoll (1997) based on evidence for Italian long-dated bonds.

8. Considering endogenously determined trading limits might be interesting, but any given limit is unlikely to change over the microstructure horizon we are considering here.

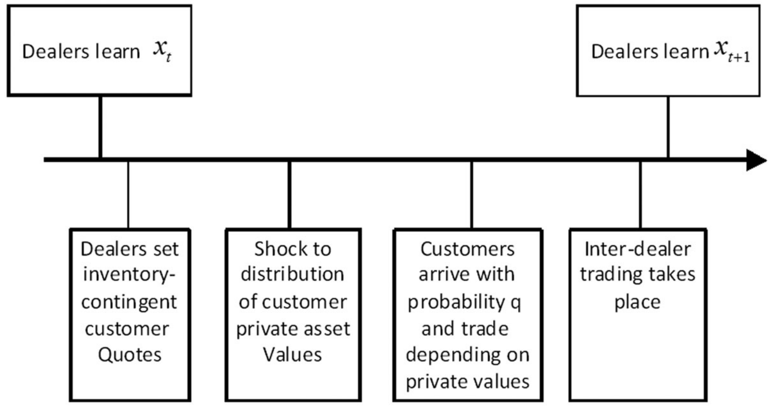

FIGURE 2. Time line for the trading process.

The matrix element $p_{12} + p_{11}$ in the first row and column arises from two possible events. Starting from a maximum inventory of 1, the dealer remains in that state if she does not conduct any trades in the B2C market: we denote this probability as p_{11}. Alternatively, the dealer might acquire an additional unit if her bid quote is accepted by a customer. In this case, the dealer would exceed the maximum inventory level of 1 and has to offset the excess inventory immediately in the B2B market with a sell transaction. We denote this probability by p_{12}. The symmetric case arises under a negative inventory level of -1, where we distinguish as p_{-1-2} the probability of a dealer selling an additional unit with the obligation to acquire immediately one unit in the B2B market.

The transition probabilities depend on the standardized state-dependent ask quotes $a(s)$ and bid quotes $b(s)$. We can now characterize the value function for the three inventory states as

$$\mathbf{V}(s, x_t) = \begin{bmatrix} V(1, x_t) \\ V(0, x_t) \\ V(-1, x_t) \end{bmatrix} = \left(\max \ \{\hat{a}(s), \hat{b}(s)\}\right) \beta \mathcal{E}_t \left[\mathbf{MV}\left(s, x_{t+1}\right) + \hat{\mathbf{\Lambda}}\right], \quad (1)$$

where $\mathcal{E}_t$ represents the expectation operator, and $\tilde{\mathbf{\Lambda}}$ denotes the period payoff given by

$$\tilde{\mathbf{\Lambda}} = \begin{bmatrix} \tilde{\Lambda}(1) \\ \tilde{\Lambda}(0) \\ \tilde{\Lambda}(-1) \end{bmatrix} = \begin{bmatrix} \left[\hat{B} - \hat{b}(1)\right] p_{12} + \hat{a}(1) p_{10} + r x_t \\ -\hat{b}(0) p_{01} + \hat{a}(0) p_{0-1} \\ -\hat{b}(-1) p_{-10} + \left[\hat{a}(-1) - \hat{A}\right] p_{-1-2} - r x_t \end{bmatrix}.$$

The payoff in state $s = 1$ includes the profit $\hat{B} - \hat{b}(1)$ if a dealer's bid quote is executed (which occurs with probability p_{12}) and the expected profit $\hat{a}(1) p_{10}$ if the ask quote is accepted by a customer. Analogous explanations apply to the other two states. The terms rx_t and $-rx_t$ capture the opportunity cost of capital for one unit of asset held (at the price x_t) as a positive or negative inventory position, respectively.

In Online Appendix A we show that the optimal quote policy can be characterized in terms of the standardized quotes $(a(s), b(s))$ and so does not depend on the level of x_t. Formally, we can characterize the dealer value function as follows.

PROPOSITION 1 (Value Function Linearity). *The value function of the dealer is linear in price and concave in inventory levels:*

$$\begin{aligned} V(1, x_{t+1}) &= V(1, x_t) + \Delta x_{t+1} &&= V - \nabla + x_{t+1} \\ V(0, x_{t+1}) &= V(0, x_t) &&= V \\ V(-1, x_{t+1}) &= V(-1, x_t) - \Delta x_{t+1} &&= V - \nabla - x_{t+1}, \end{aligned} \tag{2}$$

where V and ∇ are two positive parameters.[9]

Proof. See Online Appendix A. □

The value function is the discounted expected cash flow from being a dealer, namely of intertemporal intermediation in the B2C market and (occasionally) using the B2B market for inventory management. For the states $s = 1$ and $s = -1$ the value function $V(s, x_{t+1})$ accounts for the momentary value of the inventory given by x_{t+1} and $-x_{t+1}$, respectively. We can also show that $V(-1, 0) = V(1, 0) < V(0, 0)$. This is intuitive, as the dealer is in a more favorable position with a zero inventory than with either extreme inventory state. A dealer with no inventory owns the two-way option of being able to absorb both ask and bid transactions in the customer segment without having to resort to the interdealer market. In the extreme inventory states, the dealer owns a one-way option. For example, with a positive inventory, a customer sell cannot be internalized and the dealer is forced into the B2B market: this reduces the value function. The parameter ∇ characterizes the concavity of the value function with respect to the inventory level. It embodies a dealer's value loss due to inventory constraints.

3.3. Optimal B2C Quotes

The first-order conditions are obtained by differentiating the value function (1) with respect to the bid and ask prices $(\hat{a}(s), \hat{b}(s))$ for each inventory state s. The first-order conditions do not depend on the price process x_t. The standardized quotes $(a(s), b(s))$ can be characterized only in terms of the interdealer spread S, the parameter ∇, and the density parameter d for the distribution of reservation prices.

For example, increasing the quoted ask price $a(1)$ in state $s = 1$ marginally by ∂a has two opposite effects. It increases the expected profit on prospective sell transactions that have a likelihood of $qF^a(R^a - x_{t+1} \geq a(1) - \Delta x_{t+1}) = q(1 - a(1)d + d\Delta x_{t+1})$ for the current period. This implies an expected profit

9. A necessary condition for existence is the usual transversality condition which requires that the present value of the future payoff be bounded.

increase of $q\,[1 - a(1)d]\,\partial a$. But a higher selling price also reduces the number of expected buyers by $(qd)\,\partial a$ and the value of each transaction is given by $a\,(1) + \nabla$. The marginal gain and loss are equalized for

$$q\,[a\,(1) + \nabla]\,d \;= q\,(1 - a(1)d)\,,$$

which implies, for the optimal ask quote,

$$a(1) = (d^{-1} - \nabla)/2.$$

Similar expressions are obtained for the two other inventory states and for the optimal bid quotes, which we summarize in the following proposition.

PROPOSITION 2 (Optimal B2C Quotes). *For every given interdealer spread* $0 < S < 2d^{-1} - 4\epsilon$ *and inventory state s, there exists a unique optimal ask and bid quote* $(a(s), b(s))$ *given by*

$$\begin{bmatrix} a\,(-1) \\ a\,(0) \\ a\,(1) \end{bmatrix} = \begin{bmatrix} \frac{1}{2d} \\ \frac{1}{2d} \\ \frac{1}{2d} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \frac{S}{2} \\ \nabla \\ -\nabla \end{bmatrix}$$

and

$$\begin{bmatrix} b\,(-1) \\ b\,(0) \\ b\,(1) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2d} \\ -\frac{1}{2d} \\ -\frac{1}{2d} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \nabla \\ -\nabla \\ -\frac{S}{2} \end{bmatrix} \tag{3}$$

which depend linearly on the concavity parameter ∇ *and the interdealer spread* S. *The value function of a dealer follows as the perpetuity value of her future expected payoffs* Λ_0 *and the expected adverse selection losses* $\mathbf{\Phi}$. *Formally,*

$$\boldsymbol{V}(s, 0) = \begin{bmatrix} V - \nabla \\ V \\ V - \nabla \end{bmatrix} = (\boldsymbol{I} - \beta \boldsymbol{M})^{-1}\left(\Lambda_0 + \mathbf{\Phi}\right). \tag{4}$$

The concavity parameter $\nabla > 0$ *is monotonically increasing in* S *and monotonically decreasing in the variance* ϵ^2 *of the midprice process* x_t.

Proof. See Online Appendix B. □

Both on the bid and ask side, the optimal B2C quotes are dispersed over a range of $0.25S + 0.5\nabla$. Realized B2C bid–ask spreads vary between the inside at $a(1) - b(-1) = d^{-1} - \nabla$ and the outside at $a(-1) - b(1) = d^{-1} + 0.5S$. The dispersion of B2C execution quality therefore increases both in the B2B spread S and concavity parameter ∇. Equation (4) implicitly defines the concavity parameter ∇ as a function of the interdealer half-spread $S/2$. A particular parameter combination

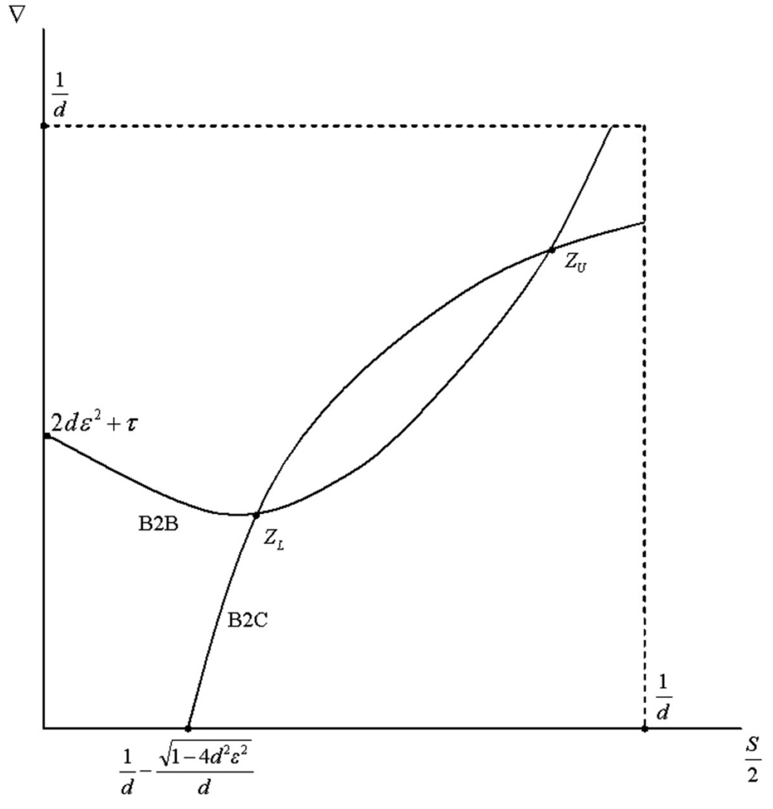

FIGURE 3. The B2C schedule characterizes the inventory concavity parameter ∇ for optimal B2C quotes under any B2B spread S. The B2B schedule defines the competitive B2B spread S for dealers who have ∇ as their inventory concavity parameter. The two intersections fulfill the equilibrium conditions in both the B2B and B2C market. Of the two equilibria, only one, Z_L, is stable.

$(S/2, \nabla)$ corresponds to optimal B2C quotes. This equilibrium schedule is graphed in Figure 3 as the B2C equilibrium schedule in a space spanned by $S/2$ and ∇. The concavity parameter ∇ monotonically increases in the B2B half-spread $S/2$. Intuitively, higher interdealer spreads render inventory imbalances more costly as rebalancing occurs at less favorable transaction prices. An increase in ∇ affects the optimal quotes differently, according to a dealer's inventory state. The optimal B2C quotes $a\,(1)$ and $b\,(-1)$ become more favorable as dealers seek to substitute B2C trades for more costly B2B trades, while B2C quotes under balanced inventories $a\,(0)$ and $b\,(0)$ deteriorate.

We can therefore conclude that a larger B2B spread S deteriorates B2C quote quality at the inventory constraints. It also magnifies the degree of inventory shading (captured by the parameter ∇) in an effort to avoid costly B2B rebalancing. The conditions $S < (2/d - 4\epsilon)$ and $\epsilon < \bar{\epsilon}$ guarantee that the optimal B2C prices fall on the support $[\pm\epsilon, d^{-1} \pm \epsilon]$ of the reservation price distribution in $t+1$. The next section develops the equilibrium condition for the interdealer market.

3.4. Competitive B2B Spreads

A competitive market structure for interdealer quotes implies that identical dealers with identical inventory levels compete away all rents in the B2B segment. Interdealer competition makes dealers indifferent as to whether their limit order is executed or not.[10] Hence, interdealer transactions do not modify the value functions of the dealers. The first-order conditions developed in Proposition 2 remain valid, even if we allow dealers to engage in B2B liquidity supply through an electronic limit order market.

Dealers with extreme inventories have a value function that is lower by $\nabla > 0$. Dealers with a negative inventory position of -1 gain ∇ by increasing their inventory level to zero and dealers with a positive inventory position *also* gain ∇ by decreasing their inventory to zero. Hence, dealers with a short inventory position will provide the most competitive interdealer bid B while dealers with a positive inventory submit the most competitive interdealer ask A. The competitive spread is therefore determined by the dealers with extreme positions who make a gross gain ∇ by moving to a zero inventory position. A larger concavity of the dealer value function with respect to inventory imbalances should (ceteris paribus) reduce the interdealer spread.

But competitive B2B limit order submission also accounts for the adverse selection risk. Limit order submission in the interdealer market also amounts to writing a trading option that other dealers can execute. In particular, we assume that a dealer with an inventory position deteriorating from -1 to -2 following a customer buy order immediately needs to rebalance to -1 by resorting to a market buy order in the interdealer market. Under Assumption 1, the distribution of the customer reservation prices is assumed to move up or down by ϵ. For example, a rise in the midprice $\left(\Delta x_{t+1} = \epsilon > 0\right)$ increases customer demand at the ask. The area of the reservation price distribution that leads to the customer acceptance of a dealer quote at the ask increases by ϵd because the reservation price distribution is uniform. This probability change is multiplied by the probability q of customer arrival to produce an upward demand shift of $\epsilon q d$. Similarly, sales at the bid to a dealer with inventory 1 fall by the same amount. Analogous remarks can be made for a fall in the midprice process.

The customer demand increase at the ask price, $a(-1)$, for a dealer with inventory -1 spills over into the B2B market. Similarly, the customer sales decrease at the bid, $b(1)$, faced by a dealer with inventory 1 is also passed on to the B2B market. The B2B market order flow is therefore correlated with Δx_{t+1}. Hence, the limit-order-submitting dealer in the B2B market is exposed to an adverse selection problem.

10. For the competitive setting to prevail, we assume that there are always (at least) two dealers with extreme positive or negative inventory positions, respectively. Bertrand competition on each side of the market then implies a competitive B2B spread.

PROPOSITION 3 (Competitive B2B Quotes). *The expected adverse selection loss due to executed limit order at both ask and bid is given by*[11]

$$L = L^A = L^B = \frac{2\epsilon^2}{(1/d) - (S/2)} > 0.$$

Under quote competition in the B2B market, the competitive ask and bid prices are given by

$$\begin{aligned} A &= \max\ \{L - \nabla + \tau, 0\} = \frac{S}{2} \\ B &= \min\ \{-L + \nabla - \tau, 0\} = -\frac{S}{2} \end{aligned} \tag{5}$$

respectively, where τ represents the order-processing costs of the liquidity provider and ∇ denotes the concavity parameter of the dealer's value function.

Proof. See Online Appendix C. □

The only occasion in which a market order is submitted is when the dealer gets pushed over the boundary from $+1$ to $+2$ or from -1 to -2. In the case of an excessive long position ($+2$), the dealer submits a sell market order while an excessive short position (-2) leads to a buy market order. A dealer that is in the $+1$ position is showing a limit sell order at the best ask in the book. Optimal B2B ask pricing ensures that the dealer is indifferent between remaining in that position or being picked off and being brought to a zero inventory status. That is why she would never submit a sell market order. Of course, a dealer in the $+1$ position would never submit a buy market order because this would unbalance her inventory.

What matters for the adverse selection loss of executed trades is not the likelihood of execution itself, but the probability of adverse midprice movement conditional on execution. The latter is not contingent on the distribution of dealers across the inventory states. Not surprisingly, the (adverse selection) loss function L is increasing in the variance ϵ^2 of the market process x_t. It is also increasing in the density d of reservation prices, because the more concentrated this distribution becomes, the greater the shift in demand induced by any given price change.

Finally, the expected adverse selection loss is increasing in the interdealer spread. Note that dealers adjust their B2C quotes $a(-1)$ and $b(1)$ to a widening B2B spread S. If B2C execution occurs nevertheless, then it is highly correlated with the directional change Δx_t of the reservation price distribution, which implies a high adverse selection risk for the liquidity suppliers in the B2B segment. Hence, adverse selection risk in the B2B market endogenously increases in the B2B spread through inventory shading in the B2C market. This feedback effect can generate market breakdown as highlighted in the Introduction: a higher S implies higher rebalancing costs and hence more price shading in the B2C market, which in turn conditions B2C execution on larger shocks to the reservation price distribution. B2B rebalancing then occurs for a more informative

11. Recall that the properties of the uniform distribution require that the denominator be positive.

customer order flow and the B2B spread S needs to increase further to reflect the higher adverse selection risk.

The equilibrium condition expressed in the second part of Proposition 3 is straightforward. A dealer with a positive inventory submits a sell limit order at the B2B ask with price A. Her expected adverse selection loss conditional on execution is L, but she gains ∇ by moving to a zero inventory if execution occurs. Under the competitive market Assumption 2, her expected conditional profit is zero, hence $A + \nabla - L - \tau = 0$, where τ represents the order-processing costs. An analogous remark applies at the bid price B. We also note that for the B2B quotes given by equation (5), dealers in inventory states $s = \pm 1$ do not find it optimal to submit market orders, as the cost $S/2$ exceeds their benefit ∇ of rebalancing. Only dealers who run against the inventory limits at ± 2 place market orders.

Proposition 3 shows that the B2B spread is given by the difference between the adverse selection loss L and the benefit of moving to a zero inventory. The interdealer quote spread is therefore negatively related to the benefit of moving to a zero inventory position and positively to the adverse selection loss of quote submission. A higher shadow cost ∇ of holding inventory imbalances therefore implies more competitive limit order submission. Very narrow B2B spreads are therefore a reflection not only of low adverse selection risk, but also of costly inventory constraints.

As with the B2C locus, we can graph the B2B locus in the $(S/2, \nabla)$ space. It is the parabola illustrated in Figure 3 with the label B2B. Its intercept and turning point are derived in the Online Appendix D.

In equilibrium, the ask price in the B2B market is the competitive price quote by a dealer on a $+1$ position who in a transaction earns his reservation price for getting back to a zero inventory (price shade) but pays the order-processing cost, the adverse selection cost, and earns the half spread. The parabolic shape is driven by $S/2$ appearing in both the adverse selection cost and the half spread. For small $S/2$ the spread earned is the dominant part and thus drives the negative relationship. For high $S/2$ the adverse selection is the largest part and thus drives the positive relationship. Dealers in the B2B market have to charge larger spreads as they realize they will attract the most "toxic" flow from the B2C market. This positive relationship between S and ∇ for high adverse selection risk is depicted by the right branch from the minimum of the parabola labeled B2B in Figure 3.

3.5. Existence and Stability of the Equilibrium

The previous sections derive separately the equilibrium relationship for the B2B and B2C markets in the $(S/2, \nabla)$ space. It is shown how the optimal quotes in the B2C market depend on the spread S in the B2B market because of rebalancing costs. Inversely, the equilibrium spread in the B2B market depends on the concavity parameter ∇ of the value function (and hence the maximum benefit of limit order submission) as well as on the degree of inventory shading which determines the degree of adverse selection of B2B market orders. This market interdependence requires that we solve the model for the joint equilibrium in both markets. The joint equilibrium solution is illustrated in Figure 3 as the intersection of the B2B and B2C graphs. Figure 3 highlights that there could be up to two equilibria. We outline why only the lower of these two equilibria is valid in the Online Appendix D.

PROPOSITION 4 (Equilibrium Existence and Stability). *Under Assumptions 1–3 and market variance ϵ^2 below some threshold $\bar{\epsilon}^2$, there exists a single stable equilibrium pair $(S/2, \nabla)$ for the B2B spread S and the concavity of the dealer value function ∇, such that (i) dealers make optimal customer quotes as stated in Proposition 2, and (ii) these quotes imply a value function with concavity ∇ so that S is the competitive B2B spread as stated in Proposition 3.*

Proof. See Online Appendix D. □

The uniqueness of the stable equilibrium Z_L allows us to undertake comparative statics with respect to the price variance ϵ^2. Note that the price volatility is directly tied to the information asymmetry between customer and dealer and the degree of adverse selection under quote provision. The axis intercepts in Figure 3 show that a variance increase (higher ϵ^2) pushes the B2B locus upwards and the B2C locus to the right. The B2B spread unambiguously increases. The same is true for an increase in the order-processing costs τ, which also shifts the B2B schedule upwards. Again, the interdealer spread S increases as the higher cost of liquidity provision in the B2B market is incorporated into the interdealer spread. But we can also highlight a small increase in order-processing costs τ—for example an exogenous security transaction tax—can induce a disproportionately larger increase in the B2B spread S. The reason here is again that higher rebalancing costs accentuate inventory shading in the B2C market and therefore increase the adverse selection risk of market orders in the B2B segment.

It is also instructive to consider two boundary cases. First, for zero volatility, the B2C schedule passes through the origin, while the intercept for the B2B curve is at the level τ. In the absence of any adverse selection, the interdealer spread reaches its minimum at a level that is less than the order-processing cost because the dealer is still partly compensated by an option value of inventory holding ∇, which remains positive. For zero order-processing costs ($\tau = 0$), the competitive interdealer spread becomes zero. Second, consider a high level of price variance given by $\epsilon^2 = 1/(4d^2)$. At this level of variance the B2C equilibrium schedule degenerates to a single point $(1/d, 0)$ without any possible intersection with the B2B locus. We conclude that at very high levels of volatility, the adverse selection effect does not allow for a market equilibrium. The market equilibrium can only exist for a volatility of the process x_t below a critical threshold so that the B2B and B2C schedules still intersect.

An interesting regulatory issue concerns the role of order-processing costs τ for market quality. The owner of a trading venue with more market power is likely to charge a higher fee for B2B transactions. Similarly, any STT on B2B trades should have the very same effect of increasing τ. In both cases the B2B schedule shifts upwards so that its intersection Z_L with the B2C schedule occurs at a higher B2B spread S

and for a higher concavity parameter ∇. More inventory concavity of the dealer value function increases the dispersion of B2C quotes. This increases the adverse selection component of the customer order flow further and reduces the incentives for market participation in the interdealer market. Thus, exogenous transaction costs can have a (negative) multiplier effect on a dealers' total incentive to engage in liquidity provision in the B2B market. Graphically, if the slopes of the B2C and the B2B schedule intersect at point Z_L, in Figure 3, at an angle of less than 45 degrees, then any small change $\Delta\tau > 0$ in transaction costs (represented by a vertical shift of the B2B schedule) will increase the interdealer half-spread $S/2$ by more than $\Delta\tau$.

Higher transaction costs also reduce market stability. In Figure 3, any upward shift of the B2B line reduces the critical level of market volatility at which market breakdown occurs. A remedy to the market destabilizing effect of higher order-processing fees is to make such fees or taxes contingent on market volatility. The optimal fee charged by the market operator should become zero or even negative when price volatility is high. This conclusion is the exact opposite of previous policy recommendations like the "Spahn tax" which propose taxation only under high levels of market volatility.

A limitation of our analysis is that dealers are risk neutral and their trading limits are exogenous; hence dealers' trading limits are assumed to be volatility invariant. In a high-volatility market, dealers might face reduced trading limits if their principals exercise active risk control. Similarly, any external change in funding liquidity for the dealer operation may also reduce inventory limits. Such additional channels for market breakdown are outside the scope of our analysis.

We note that a key role for funding liquidity in the market breakdown should imply market contagions as dealers often provide liquidity across different European sovereign bond markets. Yet, recent empirical work by Caporin et al. (2014) finds no evidence for such contagion effects in the European sovereign bond market.

4. Comparison with a Pure Retail Market Structure

We now characterize the equilibrium if dealers do not have any access to the interdealer market for rebalancing. In particular, such a pure retail market structure may arise after the interdealer market has broken down because of excessive adverse selection risk. In this case dealers may continue to engage in retail transactions, but no longer dispose of the rebalancing option of the interdealer market. Thus dealers rely exclusively on customer transactions for their inventory management. We assume that the inventory constraints are the same as in the baseline model. The solution is detailed in Online Appendix E.

The optimal quotes $a(0), b(0), a(1)$ and $b(-1)$ in equation (3) have the same functional form, but the concavity parameter ∇ generally changes and the quotes $a(-1)$ and $b(1)$ are no longer feasible, because the dealer cannot avoid the excessive inventory by immediate interdealer rebalancing. For the same reason, the previous transition probabilities p_{12} and p_{-1-2} associated with rebalancing in the upper-tier B2B interdealer market are now zero by definition. The solution involves the new concavity parameter ∇' characterized by the quadratic equation

$$f_{b2c}\left(\nabla', \epsilon^2, q, d, \beta\right) = \frac{q}{4d}\frac{\beta}{1-\beta}\left\{\left(4d^2\epsilon^2\right) - \left(1 - d\nabla'\right)^2\right\} + \nabla' = 0. \qquad (6)$$

The value maximum of the dealers' value function corresponds to the negative root

$$\nabla' = \frac{1}{2d^2 q\beta}\left[\begin{array}{c} 2d\left\{2(1-\beta) + 3q\beta\right\} - \\ \sqrt{\left\langle 2d\left\{2(1-\beta)+3q\beta\right\}\right\rangle^2 - 4d^2 q\beta\left(q\beta - 4d^2 q\beta\epsilon^2\right)} \end{array}\right], \qquad (7)$$

which exists for $\nabla' > 0$ if and only if $\epsilon^2 < 1/(4d^2)$. The volatility threshold $\bar{\epsilon}^2 = 1/(4d^2)$ marks the value of fundamental volatility at which the dealer value function becomes negative so that she would no longer provide retail quotes. It is clear from Figure 3 that in the two-tier model, market breakdown occurs at a lower volatility level, where the B2C and B2B curves are tangential. As a consequence, the upper-tier market segment is much less resilient to adverse selection and may break down even though the remaining customer–dealer tier continues to function in a restrained manner without the possibility of interdealer rebalancing.

It is also interesting to compare welfare under the two-tier market structure to that obtained under a pure retail market setting. Our comparison here focuses on the aggregate customer rents for the $2N$ potential retail customers. There are three possible dealer inventory states ($s = 1, 0, -1$), and welfare is calculated for each state as the product of the probability P_s that a trade takes place times the expected customer surplus Γ_s, so that aggregate customer welfare is defined as

$$W = 2N \sum_{s=1,0,-1} \Gamma_s P_s.$$

The probability P_s follows as the product of (i) the probability of customer arrival q, (ii) the probability p_s of the dealer inventory state s, and (iii) the likelihood of quote acceptance in each inventory state (given by $1/d$ minus the state-specific quote times d). We note that symmetry of the transition matrix implies $p_1 = p_{-1} = (1 - p_0)/2$.

It is straightforward to show (see Online Appendix E) that customer welfare in the two-tier market follows as

$$\mathbf{W}^{\substack{\text{two}\\ \text{tier}}} = Nqd\left[\begin{array}{c} p_1\left\{\frac{1}{d} - \left(\frac{1}{2d} - \frac{\nabla}{2}\right)\right\}^2 + p_0\left\{\frac{1}{d} - \left(\frac{1}{2d} + \frac{\nabla}{2}\right)\right\}^2 \\ + p_1\left\{\frac{1}{d} - \left(\frac{1}{2d} + \frac{S}{4}\right)\right\}^2 \end{array}\right], \qquad (8)$$

whereas it is

$$\mathbf{W}^{\substack{\text{retail}\\\text{only}}} = Nqd\left[p_1\left\{\frac{1}{d}-\left(\frac{1}{2d}-\frac{\nabla'}{2}\right)\right\}^2 + p_0\left\{\frac{1}{d}-\left(\frac{1}{2d}+\frac{\nabla'}{2}\right)\right\}^2\right] \quad (9)$$

in the market structure with retail trading only. We highlight that the dealer inventory state probabilities p_0 and p_1 are the same in equations (8) and (9) because the transition matrix **M** is the same under both market structures.

PROPOSITION 5 (Customer Welfare Benefits of the Interdealer Market). *A two-tiered market structure based on dealer intermediation yields higher consumer rents relative to a market structure based on retail trading only, in which rents (expressed in percent) reach only* $100 \times W^{\substack{\text{retail}\\\text{only}}} / W^{\substack{\text{two}\\\text{tier}}}$.

Under inventory constraints, customer rents are large if dealers can rebalance easily in the interdealer market, because they can offer two-sided customer quotes in all three inventory states instead of only two. The dealers' value function is more concave under pure retail trading ($\nabla < \nabla'$); yet the implied difference in retail quote quality is only of minor importance for customer welfare and does not compensate the welfare shortfall of obtaining no retail quote $a\,(-1)$ and $b\,(1)$ if the dealer is inventory constrained.[12]

A more comprehensive welfare analysis would seek to compare the two-tier market structure with a one-tier market in which retail investors can directly trade with all dealers through either limit or market orders. While we highlight the desirability of such an analysis, we found the one-tier market problem much less tractable and have to leave such work to further theoretical development. Yet a welfare comparison between these two market structures is certainly of the greatest policy relevance for the future of OTC markets.

5. Incorporating Dealer Competition

5.1. Sophisticated Customer

The model so far has assumed maximal (monopolistic) market power of dealers towards all their clients. The following section relaxes this assumption. We now assume that a share λ of te total pool of customers is highly sophisticated and able to make reverse offers to dealers. Sophisticated customers only transact if they get the best possible transaction price or do not transact at all. They offer to pay a small $\varepsilon > 0$ improvement over the shadow asset value of a dealer with an extreme inventory state, namely $x_t - \nabla + \varepsilon$ and $x_t + \nabla - \varepsilon$ on the ask and bid side, respectively. These customers will therefore obtain the best possible deals when facing a constrained dealer and extract all the rents (except ε) for themselves.

12. As no closed-form solution is available for ∇, we undertook a numerical simulation to confirm that customer rents fall to 65% from 85% if interdealer rebalancing is suspended.

ASSUMPTION 4 (Different Customer Types). A share $0 < \lambda < 1$ of customers engage in reverse offers to their dealer at retail prices $x_t - \nabla + \varepsilon$ and $x_t + \nabla - \varepsilon$ on the ask and bid side, respectively. These reverse offers represent a small $\varepsilon > 0$ improvement over the reservation price of a dealer in inventory state $s = 1$ and $s = -1$, respectively. All other customers trade as before.

An increasing share of sophisticated customers limits the overall rents a dealer can extract from his customer pool. We can also think of the sophisticated customers as those who search for the best retail deal available. Their presence proxies for a reduced-form assumption about interdealer competition.

We highlight that this model extension is relatively tractable because additional transactions at dealer reservation prices only alter the transition probabilities between inventory states in the matrix $\mathbf{M}$. While this changes the value function and its concavity parameter ∇, this does not alter—except indirectly through ∇—the first-order conditions in Proposition 2. Moreover, the introduction of the sophisticated traders does not affect the B2B equilibrium schedule in Figure 3. Only the B2C schedule undergoes a shift as shown in the Online Appendix F.

5.2. Equilibrium Effects of Dealer Competition

The consequences of dealer competition can be summarized by the following proposition.

PROPOSITION 6 (Market Equilibrium with Sophisticated Customers). *A larger share λ of sophisticated customers shifts the B2C schedule in Figure 3 downwards with the following implications:*

(1) *the interdealer spread S monotonically increases;*
(2) *the concavity parameter ∇ decreases (increases) at low (high) volatility;*
(3) *more retail transactions (by sophisticated customers) occur inside the interdealer spread;*
(4) *market breakdown occurs at lower volatility (and less adverse selection).*

Proof. See Online Appendix F. □

The intuition for these results is relatively simple. Sophisticated customers provide dealers with an additional rebalancing opportunity captured by the relative increase of the off-diagonal elements $(1, 2)$ and $(3, 2)$ in a revised transition matrix $\mathbf{M}_\lambda$. This increases the likelihood of a balanced inventory state and makes the dealers less likely to resort to costly rebalancing in the interdealer market. For any given rebalancing cost given by the interdealer spread S, the concavity parameter ∇ should therefore take on a lower value. But this exactly corresponds to a downward shift in the B2C schedule. Implications (1)–(4) then directly follow from the graphical analysis provided in Figure 3 for an unchanged B2B schedule. Under a downward B2C shift, the stable equilibrium point Z_L moves to the right, which corresponds to a higher interdealer market spread S. If the stable equilibrium Z_L is situated on the left (right) branch of the B2B schedule, then the downward shift of the B2C schedule decreases (increases) the concavity parameter ∇; implying less (more) price discrimination across inventory states and therefore more B2C price dispersion among nonsophisticated customers. Finally, a flatter B2C schedule makes it more likely that no intersection with the B2B schedule occurs; hence the higher market fragility of the dual market structure under an increased share of sophisticated rent-capturing customers. Because of the reduction in the market power of dealers, the novel empirical claim of this paper is reinforced: customer–dealer spreads are more likely to be smaller than interdealer spreads.

6. Evidence from the European Sovereign Bond Market

The market structure in the European sovereign bond market corresponds to the two-tier framework captured in our model of dealer intermediation. The following empirical analysis focuses on three key predictions of our model; namely (i) a large dispersion of B2C spreads due to inventory-contingent dealer pricing; (ii) an increase of both the B2B spread and the B2C price dispersion in bond maturity as a reflection of adverse selection risk; and (iii) evidence that the dealers' aggregate inventory situation (as proxied by B2B limit order imbalances) directly correlates with the quality of B2C ask- and bid-side transactions with the opposite sign.

6.1. Market Overview

The market participants in the European bond market can be grouped into primary dealers, other dealers, and customers. Customers are typically other financial institutions, like smaller banks or investment funds. Dealers have access to electronic interdealer (B2B) platforms, of which the most important is MTS. Its largest market share is in Italy, where it has close to 100%. In other countries MTS has a lower market share but overall, approximately half of all interdealer trades are transacted through MTS.[13] Trading in the MTS interdealer platform is similar in operation to any electronic limit order book market.

At the time of our study, B2C transactions took place both over-the-counter and on various trading platforms. The Eurex platform had not long been established and did not have a large share of the market. Also, Bloomberg's BBT platform was mostly a repository for limit orders and expressions of interest in awkwardly sized or very small orders. TradeWeb and BondVision customers were able to submit simultaneously "requests-for-quotes" (RFQs) from a small number of dealers who could potentially supply instant responses that could be accepted electronically. Though TradeWeb has a slightly larger market share than BondVision, the latter is operated by MTS in parallel with its B2B platform and thus it was easier to compile consistent and accurate time-stamped data from the two segments by using BondVision data. The BondVision

13. For more institutional background, see also Dunne, Moore, and Portes (2006, 2007).

platform represents a significant proportion of B2C electronic RFQ trading, particularly for Italian issues. Given the strong market position of MTS in the Italian B2C segment, it is natural to focus much of our empirical analysis on Italian bonds.

6.2. *MTS and BondVision Data*

We explore a new data set that combines both interdealer (B2B) and dealer–customer data (B2C). The data cover the last three quarters of 2005. Events are reliably time stamped and trade initiation is electronically signed in both markets. In the case of the B2B market we obtained observations about the state of the limit order book at a per second frequency and we were also provided with transaction data on an event basis. Our empirical analysis involves a comparison of the transactions with customers on the BondVision platform with the prevailing quotes made between dealers on the B2B platform at the exact time of the customer requests for quotes.

Over the data period 72 (resp. 268) different Italian (non-Italian) bonds were traded on both MTS and BondVision. Our sample consists of 105,469 (83,313) Italian (non-Italian) bond B2B trades and 28,245 (resp. 17,259) Italian (non-Italian) bond B2C trades. The majority of trades in each case concern so-called benchmark bonds.[14] The "benchmark" attribute that we employ is defined by MTS and refers to bonds for which primary dealers have liquidity provision obligations. We also group the bonds into three different maturity groups. Short–medium bonds have a maturity of 1.5–7.5 years, long bonds of 7.5–13.5 years, and very long bonds feature maturities beyond 13.5 years.

The unique feature of our data is that they combine interdealer and dealer–customer price data. It is therefore straightforward to assess the competitiveness of the B2C segment by comparing the B2C trades to the best B2B quote at the same side of the market. We distinguish B2C trades that occur at the ask and compare them to the best B2B ask price prevailing at the same moment in time. Similarly, B2C trades at the bid side of the market are compared to the best available contemporaneous B2B bid price. We refer to this price difference as cross-market spread, defined as

$$\begin{aligned}\text{Cross-Market Spread (Ask)} &= \\ \text{Best B2B Ask Price} &- \text{B2C Ask Price} \\ \text{Cross-Market Spread (Bid)} &= \\ \text{B2C Bid Price} &- \text{Best B2B Bid Price.}\end{aligned}$$

We present three strands of evidence. Firstly we provide a nonparametric analysis of cross-market spreads under different categories of bond liquidity. Secondly, we provide a similar analysis across different bond maturities. Finally, we carry out a full regression analysis to test the model implications for market volatility.

14. For more institutional background, see also Dunne, Moore, and Portes (2006, 2007).

6.3. Market Quality by Bond Liquidity

How favorable are B2C transaction prices in BondVision relative to the best B2B quote on the same side of the market in the MTS interdealer platform? Table 1 addresses this question for the total sample of 340 bonds. It reports the cross-market spread for ask-side trades and (separately) bid-side trades for bonds in the four liquidity groups. The four liquidity categories are a two-by-two classification by Italian/non-Italian and benchmark/nonbenchmark bonds. The cross-market spreads for each liquidity category are grouped into quartiles, where Q(1) denotes the 25% lowest (best) cross-market spreads and Q(4) represents the 25% highest (worst) spreads from the customer perspective. We report the quartile mean as well as the overall mean.

The insight from Table 1 concerns both the overall quality of B2C trades as well as their large dispersion relative to the best B2B quotes. First, the average B2C trade quality appears high. The mean cross-market spread is positive for Italian and non-Italian bonds, for benchmark and nonbenchmark bonds and on both bid- and ask-side transactions. Even the mean of the 25% worst B2C transactions on the ask side shows a slightly positive cross-market spread. Their execution quality is therefore too favorable relative to the negative cross-market spread $0.5S - a(-1) < 0$ predicted by the model for the worst B2C transactions. On the bid side, B2C trades are slightly less favorable. The 25% worst trades show an average transaction price outside the B2B spread in line with the model prediction. The cross-market spread is somewhat smaller for Italian benchmark bonds compared to the other three categories. But the overall finding is similar across all four groups. B2C transactions occur on average at or inside the B2B spread. Second, the dispersion of the cross-market spread is substantial. It ranges from an average of 4.80 (4.75) cents for the 25% best B2C ask (bid) side trades to 0.24 (−0.38) cents for the 25 worst B2C ask (bid) side trades. This is large relative to an average interdealer (B2B) spread of approximately 4.31 cents. It is our contention that such quality dispersion of B2C trades can be explained by our model of inventory-contingent dealer quotes.

The right-hand side of panels A and B report the distribution of B2B spreads recorded at the time when B2C trades occur. On the ask side, the average B2B half-spread is 1.98 cents (≈1.98 basis points) and can be compared to the average cross-market spread of 1.99 cents (≈1.99 basis points). This implies that ask-side B2C trades occur on average at the midpoint of the B2B spread. On the bid side, B2C trades are slightly less favorable, but still extremely “low cost”. B2C trades are centered around a price level between the B2B midprice and the best B2B bid price, as the comparison between the average cross-market spread of 1.49 cents and the B2B half-spread of 2.33 cents reveals.

6.4. Market Quality by Bond Maturity

One explanation for the large dispersion of B2C trade quality is dealer price discrimination by customer type. Less sophisticated customers may for example obtain

TABLE 1. Cross-market spreads and B2B spreads by liquidity.

Panel A: Ask-side spreads*																
		B2B ask spreads $A - MidP$					B2C ask spreads $a - MidP$					Cross-market ask spreads $A - a$				
		Italian bonds		Non-Italian bonds			Italian bonds		Non-Italian bonds			Italian bonds		Non-Italian bonds		
Quantile means	Quality	B	NB	B	NB	All	B	NB	B	NB	All	B	NB	B	NB	All
Mean of $Q(1)$	Best	0.64	0.24	0.90	0.89	0.70	−1.20	−1.65	−1.91	−1.66	−1.56	3.82	5.90	5.58	5.22	4.80
Mean of $Q(2)$		1.00	0.62	1.38	1.16	1.04	−0.26	−0.29	−0.42	−0.48	−0.35	1.56	1.44	2.00	2.00	1.93
Mean of $Q(3)$		1.29	1.20	1.66	1.50	1.52	0.08	0.02	0.15	0.00	0.09	1.00	0.91	1.37	1.28	1.00
Mean of $Q(4)$	Worst	3.86	5.93	5.10	4.88	4.64	1.34	1.45	2.27	1.72	1.74	0.45	0.20	0.01	0.35	0.24
Overall mean		1.70	1.99	2.26	2.11	1.98	−0.01	−0.12	0.02	−0.11	−0.02	1.71	2.11	2.24	2.21	1.99

Panel B: Bid-side spreads																
		B2B bid spreads $MidP - B$					B2C bid spreads $MidP - b$					Cross-market bid spreads $b - B$				
		Italian bonds		Non-Italian bonds			Italian bonds		Non-Italian bonds			Italian bonds		Non-Italian bonds		
Quantile means	Quality	B	NB	B	NB	All	B	NB	B	NB	All	B	NB	B	NB	All
Mean of $Q(1)$	Best	0.67	0.60	0.91	0.89	0.76	−0.34	−1.03	−1.61	−0.82	−0.94	3.13	6.42	5.26	4.27	4.75
Mean of $Q(2)$		1.00	2.84	1.40	1.16	1.17	0.55	0.83	0.45	0.39	0.52	1.00	3.23	1.19	1.00	1.16
Mean of $Q(3)$		1.46	5.43	1.69	1.50	1.74	1.00	1.78	0.92	0.79	1.05	0.00	1.14	0.72	0.68	0.43
Mean of $Q(4)$	Worst	4.14	7.13	5.17	4.68	5.66	2.15	3.95	2.79	2.55	2.75	−0.22	−0.32	−0.56	−0.59	−0.38
Overall mean		1.82	4.00	2.29	2.06	2.33	0.84	1.38	0.64	0.73	0.84	0.98	2.62	1.65	1.33	1.49

Notes: We report for each quantile of the trade price distribution (i) the average B2B spread, (ii) the average B2C spread and (iii) the average of the cross-market spread for 72 Italian and 268 non-Italian European sovereign bonds of high (B = benchmark) and low (NB = nonbenchmark) liquidity. Panel A reports average spreads for transactions at the ask quotes while Panel B reports spreads for bid transactions. The B2B or B2C spreads are measured relative to the midprice *MidP* between the best B2B ask and bid at the same moment in time when the B2B or B2C transactions occur. The cross-market spread is defined as the difference between the B2C transaction price (a or b for B2C ask or bid, respectively) and the prevailing best B2B price (A or B for B2B ask or bid, respectively). All spread measures are given in cents. At par, these amount to basis points.

systematically worse B2C quotes. Under this alternative hypothesis, the B2C price dispersion should be unrelated to the adverse selection risk and inventory constraints of the dealers. While we cannot sort cross-market spreads by customer type (for lack of customer information), we can reproduce Table 1 sorted by bond maturity. Long-run bonds have a higher duration and their larger interest rate sensitivity implies that price volatility and adverse selection risk are considerably larger than for bonds of short maturity. According to our model of inventory-based price differentiation, the B2C price dispersion increases in midprice volatility and therefore also in bond maturity.

Table 2 presents cross-market spreads for 171 benchmark bonds (Italian and non-Italian) classified by three maturity groups. The mean B2B ask (bid) side half-spread in Panel A (Panel B) increases from 1.01 (0.99) cents to 5.38 (5.12) cents when comparing very long bonds to short–medium bonds. This fivefold increase highlights the strong sensitivity of the B2B to adverse selection risk. By contrast, the mean cross-market spread shown on the right-hand side of Table 2 increases less on both the ask and bid side, which implies higher relative transaction quality for B2C transaction as monopolistic dealers absorb some of the adverse selection risk in the B2C segment. The dispersion of the cross-market spread between the 25% best and worst B2C trades is 1.75 (1.48) cents on the ask (bid) side for short and medium maturities and increases to 9.59 (9.21) cents on the ask (bid) side for the very long maturities. The B2C price dispersion therefore increases by more that a factor of five for bonds of high duration. This feature of the data cannot be accounted for by customer-based price discrimination since customers of very different financial sophistication are likely to request both long and short maturity bonds. Overall, the data sort on bond maturity suggests that B2C trade quality dispersion is driven by a dealer's inventory management costs (i.e., the cost of rebalancing in the B2B market) rather than a pure customer-based price discrimination.

6.5. Market Quality by Inventory Imbalances and Market Volatility

It is clear from Figure 4 that an implication of the model is that higher adverse selection, as measured by volatility, implies that the quality of the average B2C spread should improve relative to the B2B spread. So the average cross-market spread should decrease in volatility on both the ask and bid sides of the market. The other important feature of the model is that the B2C quotes depend on the inventory state of the dealer. Unfortunately, such inventory data are not directly available. However, inventory imbalances also induce dealers to submit the most competitive B2B quotes. The relative depth of the best B2B quotes indicates the distribution of inventory imbalances within the dealer population. Therefore we measure aggregate inventory imbalances as

$$Imb = \frac{Q(Ask) - Q(Bid)}{Q(Bid) + Q(Ask)},$$

where $Q(\cdot)$ denotes the limit order book liquidity at the best ask or bid, respectively.

TABLE 2. Cross-market spreads and B2B spreads by bond maturity.

Panel A: Ask-side spreads*

		B2B ask spreads $A - MidP$				B2C ask spreads $a - MidP$				Cross-market ask spreads $A - a$			
		Bond maturity				Bond maturity				Bond maturity			
Quantile means	Quality	Short–med.	Long	Very long	All	Short–med.	Long	Very long	All	Short–med.	Long	Very long	All
Mean of $Q(1)$	Best	0.52	0.96	2.53	0.76	−0.98	−1.64	−2.73	−1.53	2.21	3.26	9.40	4.64
Mean of $Q(2)$		0.99	1.39	4.68	1.06	−0.37	−0.43	0.13	−0.34	1.16	2.00	5.20	1.94
Mean of $Q(3)$		1.00	1.50	6.02	1.53	0.00	0.11	1.33	0.11	1.00	1.17	3.35	1.00
Mean of $Q(4)$	Worst	1.54	2.23	8.28	4.48	0.56	1.21	5.01	1.77	0.46	0.39	−0.19	0.24
Overall mean		1.01	1.52	5.38	1.96	−0.20	−0.19	0.93	0.00	1.21	1.71	4.44	1.95

Panel B: Bid-side spreads

		B2B bid spreads $MidP - B$				B2B bid spreads $MidP - b$				Cross-market bid spreads $b - B$			
		Bond maturity				Bond maturity				Bond maturity			
Quantile means	Quality	Short–med.	Long	Very long	All	Short–med.	Long	Very long	All	Short–med.	Long	Very long	All
Mean of $Q(1)$	Best	0.49	0.95	2.25	0.78	−0.16	−0.83	−2.53	−0.92	1.23	2.51	9.10	4.17
Mean of $Q(2)$		0.97	1.36	4.33	1.14	0.48	0.48	0.94	0.49	0.54	1.00	4.10	1.00
Mean of $Q(3)$		1.00	1.50	5.72	1.61	0.84	0.96	1.97	0.99	0.00	0.46	2.36	0.32
Mean of $Q(4)$	Worst	1.51	2.37	8.18	4.60	1.29	1.96	4.66	2.44	−0.25	−0.36	−0.11	−0.37
Overall mean		0.99	1.55	5.12	2.03	0.61	0.64	1.26	0.75	0.38	0.90	3.86	1.28

Notes: We report for each quantile of the trade price distribution (i) the average B2B spread, (ii) the average B2C spread and (iii) the average cross-market spread for a sample of bonds grouped into three main maturity categories of 171 (Italian and non-Italian) benchmark bonds. Panel A reports average spreads for transactions at the ask quotes while Panel B reports spreads for bid transactions. The B2B or B2C spreads are measured relative to the midprice *MidP* between the best B2B ask and bid at the same moment in time when the B2B or B2C transactions occur. The cross-market spread is defined as the difference between the B2C transaction price (a or b for B2C ask or bid, respectively) and the prevailing best B2B price (A or B for B2B ask or bid, respectively). All spread measures are given in cents. At par, these amount to basis points.

Figure 4, panel A plots the average cross-market spread $A - \bar{a}$ on the ask side as a function of the inventory imbalance and the volatility. The corresponding cross-market spread $\bar{b} - B$ on the bid side is featured in panel B. As before, higher volatility increases this spread because of the higher volatility sensitivity of the B2B spread S. Moreover, Figure 4 also reveals the dependence of the cross-market spread on the inventory imbalance. A more positive aggregate inventory imbalance, namely more dealers in state $s = 1$ relative to $s = -1$, comes with a lower average ask quote $\bar{a}$ and therefore a higher cross-market spread on the ask side. On the bid side, the cross-market spread decreases in the imbalance statistic, as depicted in panel B. Intuitively, a positive imbalance comes with a tilt of the probability distribution of dealer states toward $s = 1$. This implies that relatively more dealers quote B2C prices $a(1)$ or $b(1)$ relative to $a(-1)$ or $b(-1)$. Hence the average cross-market spread improves on the ask side and deteriorates on the bid side. The previous regression is now extended as follows:

$$\begin{aligned}
&\text{Cross-Market Spread (Ask);}\\
&\qquad A - a = \mu_{a0} \ + \ \mu_{av} \times Vol + \mu_{aI} \times Imb + \eta_a\\
&\text{Cross-Market Spread (Bid);}\\
&\qquad b - B = \mu_{b0} \ + \ \mu_{bv} \times Vol + \mu_{bI} \times Imb + \eta_b,
\end{aligned}$$

where η_a and η_b are i.i.d. processes, and μ_{a0}, μ_{av}, μ_{aI}, μ_{b0}, μ_{bv}, and μ_{bI} are parameters. The null hypotheses are that $\mu_{av} = \mu_{bv} > 0$ and $\mu_{aI} = -\mu_{bI} > 0$.

A potential problem with this regression is simultaneity bias. Price outliers in the interdealer market tend to influence both the B2B half-spread and the volatility measurement in the same period. To avoid this simultaneity bias, we use again an instrumental variable approach based on lagged rather than contemporaneous volatility. We also include fixed effects for each bond to control for heterogeneity across bonds.

In Table 3, columns (10) and (12) present the regression results for the cross-market spread. Panel A reports the regression results for the ask side and panel B for the bid side of the market. The analysis here focuses on the Italian bonds because of the high market coverage of our B2C data for this segment. In each case we run a regression for the full sample of all 13 liquid Italian government bonds and the subsample of six most liquid long-dated Italian government bonds. The six long-dated bonds form a particularly homogeneous subsample in terms of coupon rates, maturity, and liquidity characteristics, and at the same time represent a large share of the overall bond transactions in Italian long-dated bonds.[15] The cross-market spread on the ask side is almost constant in volatility and increasing on the bid side. The increase on the bid side is statistically significant at the 1% level for the full sample though the significance is marginal for the subsample of long maturity bonds. For the ask side, we cannot confirm that the predicted cross-market spread increases in volatility. Hence,

15. The results are also conditioned on two controls. The log of B2C transaction size controls for trade size while competition effects are controlled for by the use of separate intercepts for RFQs from a single dealer and RFQs from more than one dealer.

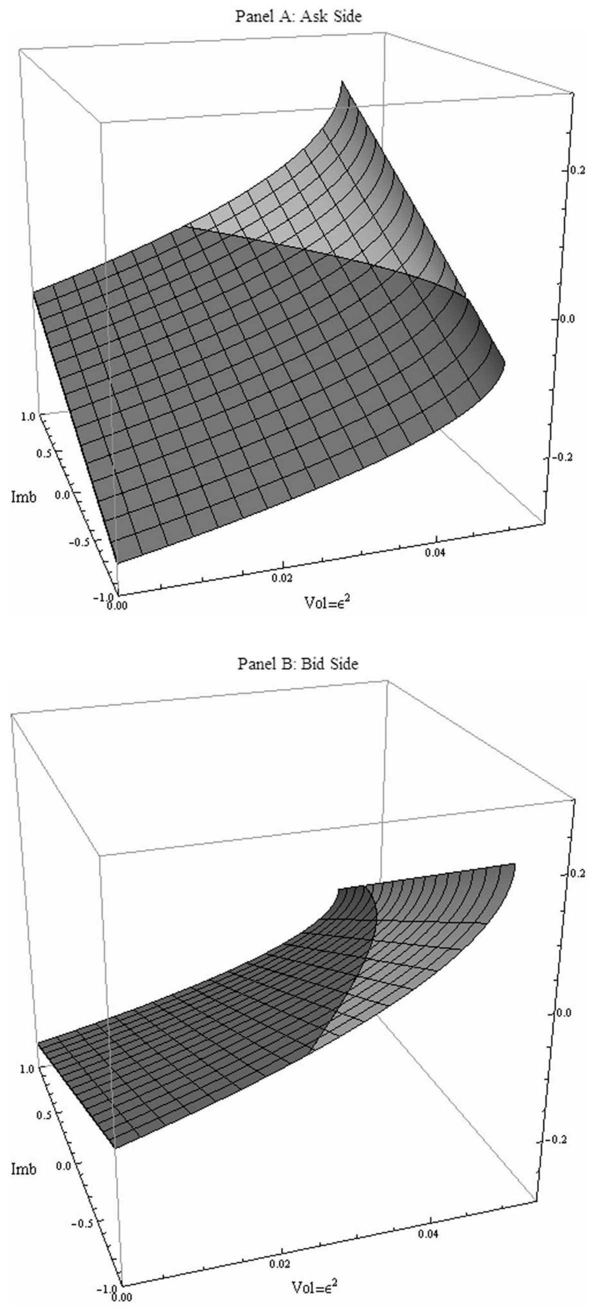

FIGURE 4. For the ask side (panel A) and the bid side (panel B) we plot vertically the average cross-market spread as a function of volatility (ϵ^2) and the aggregate inventory imbalance (*Imb*). The darker area marks the region for which the average B2C spread is more favorable than the B2B spread. The order-processing cost parameter is chosen as $\tau = 0.5$; the probability of customer arrival is $q = 0.5$; the discount rate is $\beta = 0.99$; the density of the customer price reservation distribution d is set at 1.

TABLE 3. Cross-market spread and B2B spread estimation.

Panel A: Ask-side spreads*

	B2B spread				B2C spread				Cross-market spread			
	Full sample		Long bonds		Full sample		Long bonds		Full sample		Long bonds	
Regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Long Realized volatility	0.277	0.277	0.408	0.406	0.291	0.293	0.329	0.355	0.01	0.013	−0.085	−0.06
T-Stat	4.719	4.714	3.130	3.126	3.504	3.544	1.836	2.009	0.24	0.305	−0.918	−0.667
Imbalances, *Imb*		−0.037		−0.040		−0.265		−0.441		0.330		0.477
T-Stat		−1.316		−0.820		−5.712		−5.273		11.85		8.724
NO COMP	−0.023	−0.025	−0.919	−0.879	0.115	0.124	0.125	−0.308	0.423	0.436	0.974	0.518
T-Stat	−0.099	−0.104	−1.4999	−1.454	0.336	0.366	0.141	−0.357	2.355	2.457	2.075	1.144
COMP 2+					0.246	0.252	0.830	0.370	0.576	0.584	1.577	1.092
T-Stat					0.749	0.769	0.991	0.452	3.348	3.444	3.627	2.599
Log B2C quantity					−0.129	−0.128	−0.234	−0.236	−0.067	−0.07	−0.11	−0.113
T-Stat					−7.188	−7.133	−6.322	−6.443	−7.012	−6.908	−4.929	−5.169
Obs.	5159	5159	1561	1561	5159	5159	1561	1561	5159	5159	1561	1561
OLS $\bar{R}^2$ (no fixed effects)	0.833	0.833	0.436	0.445	0.802	0.803	0.318	0.325	0.561	0.570	0.061	0.096
$F(3)$					1.974	1.901	3.684	3.411	3.650	3.387	4.906	4.501

TABLE 3. Continued.

Panel B: Bid-side spreads												
	B2B spread				B2C spread				Cross-market spread			
	Full sample		Long bonds		Full sample		Long bonds		Full sample		Long bonds	
Regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Long Realized volatility	0.554	0.555	0.590	0.590	0.707	0.698	0.703	0.700	0.15	0.143	0.122	0.119
T-Stat	8.020	8.026	5.536	5.547	7.033	6.961	4.459	4.432	2.764	2.638	1.568	1.535
Imbalances, *Imb*		0.019		0.074		0.290		0.278		−0.313		−0.345
T-Stat		0.537		1.763		5.213		3.931		−9.757		−7.545
NO COMP	−1.011	−1.015	2.317	2.319	−1.583	−1.536	3.040	3.005	0.469	−0.425	0.921	0.878
T-Stat	−3.673	−3.684	5.628	5.642	−3.807	−3.699	4.496	4.449	−2.105	−1.904	2.462	2.386
COMP					−1.414	−1.406	2.996	2.993	−0.468	−0.406	0.641	0.637
T-Stat					−3.450	−3.531	4.787	4.776	−2.2	−1.907	1.972	1.969
Log B2C quantity					−0.086	−0.082	−0.136	−0.135	−0.063	−0.059	−0.074	−0.073
T-Stat					−4.385	−4.213	−3.863	−3.845	−5.47	−5.18	−3.565	−3.529
Obs	4441	4441	2082	2082	4441	4441	2082	2082	4441	4441	2082	2082
OLS $\bar{R}^2$ (no fixed effects)	0.820	0.820	0.432	0.435	0.782	0.783	0.318	0.321	0.517	0.524	0.076	0.091
$F(3)$					0.718	2.640	0.277	0.223	2.138	1.806	5.601	4.972

Notes: Reported are instrumental variable estimates of the relation between the spreads, volatility, and imbalance controlling for competition and order size where applicable. The dependent variables are (i) the B2B spread (columns (1)–(4)), (ii) the B2C spread (columns (5)–(8)), and (iii) the cross-market spread (columns (9)–(12)) for the ask side (Panel A) and the bid side (Panel B), respectively. The explanatory variables are realized volatility and imbalance at the best quotes in the B2B market prevailing at the time of the B2C request for quotes. Volatility is measured by the log-realized volatility of the midprice returns over one-minute intervals computed for every full hour. Imbalance (*Imb*) is measured as the difference between the B2B liquidity at the best ask and the best bid for the benchmark Italian long bond at the moment when a B2C transaction takes place in any given bond. The competition control is in the form of separate dummies for requests for quotes from one dealer and more than one dealer respectively. Order size enters as the log of B2C quantity. Results are provided for the full-sample of liquid Italian bonds and for the subsample containing the six very liquid long bonds. In all cases we include bond-specific fixed effects to control for spread differences across bonds. The IV regression uses a constant and volatility lagged by one hour as instruments. The t-statistics presented are based on standard errors that have been adjusted for heteroskedasticity. Spreads are expressed in cents. At par, these amount to basis points. Even-numbered regressions include the imbalance variable. The reported $\bar{R}^2$ are from OLS regressions with fixed effects: they are higher for the full sample regressions because of the much larger number of bonds. The F-tests are for equality of the constants for competition/no-competition in regressions (5)–(12).

there is no change in the B2C ask-side trade quality (relative to the best B2B quote) as volatility changes.

The results for the inventory dependence of the cross-market spread are more clear-cut. The estimation coefficients have the signs predicted under the null hypothesis and are therefore consistent with the numerical results depicted in Figure 4.[16] The imbalance measure itself is statistically highly significant with t-statistics always above 7 in absolute value. For the ask side we find a positive effect on the cross-market spread and for the bid side a negative coefficient as proposed under the null.

The B2B spreads in Table 3, columns (1)–(4), show, as expected, a highly significant positive volatility dependence. The volatility dependence in the full sample is stronger on the bid side than the ask side with coefficients 0.554 and 0.277, respectively. The more positive volatility dependence for the B2B spread on the bid side may explain algebraically why we find a more positive volatility dependence for the cross-market spread on the bid side as well. The asymmetry in the spread behavior between the ask and bid side needs to be explained by forces outside the current model framework. It is reassuring that the B2B spreads do not display a pattern of statistically significant dependence on inventory imbalances. For completeness, columns (5) to (8) of Table 3 display the results of using B2C spreads as the dependent variable

Finally, we highlight that the point estimates, in absolute value, for imbalances in the cross-market spread equations in columns (9)–(12) of Table 3 vary between 0.313 and 0.477: these are also economically significant. To see this, assume that inventory imbalances move over half the maximal range from -0.5 to 0.5. The coefficient estimates then represent the corresponding change in the B2C price quality in cents. Such an inventory-related price change is large considering that, as Table 3 shows, the B2B half-spreads are on average only 1.40 cents on the ask side and 1.68 cents on the bid side whenever B2C trades occur. A two standard deviation increase in the imbalance variable improves ask-side B2C transactions by 0.42 basis points and deteriorates bid-side transactions by 0.30 basis points. Inventory imbalances proxied by liquidity imbalances in the B2B market therefore explain economically significant variations in B2C transaction price quality.

7. Extensions and Limitations of the Analysis

Our simple dynamic market intermediation problem of optimal B2B and B2C price setting already gives rise to a relatively rich model in the case of only three inventory states. Here we point out some possible extensions.

A first generalization is to extend the number of inventory states from 3 to $2n + 1$. Since every inventory state comes with separate first-order conditions for the B2B and B2C segment, we would have to solve $4n + 2$ equations. Instead of a single concavity

16. The imbalance measure is almost orthogonal to the volatility measure (their correlation is a mere 0.0076) and its inclusion in the regression is without consequence for the spread–volatility nexus as is clear from the odd-numbered columns.

parameter ∇, we would have to solve for a set of n value function parameters. But we do not see that this increased complexity renders any new qualitative insights into the dynamics of the intermediation problem.

A second more interesting extension consists of allowing for asymmetry of the reservation price distribution on the ask and bid side. Summary statistics in Tables 1 and 2 show somewhat more favorable cross-market spreads on the ask than on the bid side. One straightforward explanation could be that the distribution of customer reservation prices is more dense on the ask side. The model can capture this by distinguishing the ask-side distribution of reservation prices by a parameter d_a from the corresponding bid-side parameter d_b with $d_a > d_b$. This symmetry-breaking assumption implies that first-order conditions on the ask and bid side are no longer mirror images and the value function is no longer symmetric in inventory imbalances. We rather obtain separate concavity parameters ∇_a and ∇_b influencing ask- and bid-side quotes differently. While this is still rather tractable and can capture bid- and ask-side asymmetry, we conjecture that the fundamental insights of the models are not altered.

A still more desirable extension would be the introduction of a more general form of dealer competition for customer quotes. The model extension in Section 6 provides a first parsimonious step towards modeling reduced dealer market power, but its stylized dichotomy between price-taking and price-setting customers is not fully satisfying. Yet, more general extensions pose fundamental challenges. Simple Bertrand price competition in a dealer duopoly already eliminates all price-setting power for the dealers. Such a fully competitive setting would be at odds with the evidence for inventory effects. In order to moderate price competition and retain some price setting power for dealers, additional assumptions are needed. It seems technically difficult to introduce a more general version of interdealer competition for customer quotes into our framework.

While our model allows for a relatively straightforward welfare analysis of the two-tier market structure, it does not inform us how this welfare compares to the one-tier market in which both traders and retail customers interact through a single limit order market. Such a comparison should be considered a high priority for future research as current regulatory policy aims at restraining (two-tier) OTC structures in favor of (one-tier) trading in centralized exchanges.

8. Conclusions

Repeated market breakdown in the European sovereign bond market during the financial crisis calls for a better understanding of adverse selection problems in a two-tier market structure. The current paper develops a theoretical framework which allows for a better understanding of dealers as intermediaries between a highly competitive centralized interdealer trading platform (B2B) and a network of client relationships (B2C). We characterize the interrelationship between both market segments and its fragile nature. First, adverse selection risk passes from the client network to the interdealer market, where it may generate market breakdown. This happens easily if the dealers' asset valuation differ (as in our model) only by rebalancing benefits. The interdealer segment functions only so long as the benefit of inventory rebalancing exceeds the cost of adverse selection. Second, the interdealer spread determines the rebalancing costs for the dealers and therefore feeds back to the degree of inventory shading, retail price dispersion, and the average retail spread. Third, if retail spreads increases, this tends to increase the adverse selection component of the client order flow, implying still higher interdealer spreads. Such a feedback loop can easily generate market breakdown in the interdealer segment.

Our analysis has important regulatory policy conclusions. Low order-processing costs in the interdealer market are important for the robustness of the market structure. This implies that the market power of the interdealer platform provider should be a prime regulatory concern. We find indeed that the interdealer spreads in the European sovereign bond trading platform MTS are large relative to the B2C spreads available outside the centralized market. This points to relatively important order-processing costs, which should make the market more fragile and susceptible to market breakdown. At the very least one would expect full public disclosure about such order-processing costs—something which is not the case today. Any increase of the order-processing costs due to STT is also detrimental to market stability, as shown in our analysis. The current regulatory debate about such taxes could benefit from the structural analysis provided in this paper.

References

Brunnermeier, Markus K. and Lasse H. Pedersen (2009). "Market Liquidity and Funding Liquidity." *Review of Financial Studies*, 22, 2201–2238.

Caporin, Massimiliano, Lorianna Pelizzon, Francesco Ravazzolo, and Roberto Rigobon (2014). "Measuring Sovereign Contagion in Europe." *SSRN*. Available at http://ssrn.com/abstract=2023756.

Dunne, Peter G., Michael J. Moore, and Richard Portes (2006). "European Government Bond Markets: Transparency, Liquidity, Efficiency." *City of London Corporation Monograph commissioned from CEPR*. Available at http://www.cityoflondon.gov.uk/economicresearch.

Dunne, Peter G., Michael J. Moore, and Richard Portes (2007). "Benchmark Status in Fixed-Income Asset Markets." *Journal of Business Finance and Accounting*, 34, 1615–1634.

Duffie, Darrell (2010). "Asset Price Dynamics with Slow-Moving Capital." (American Finance Association Presidential Address), *Journal of Finance*, 65, 1238–1268.

Duffie, Darrell and Bruno Strulovici (2012). "Capital Mobility and Asset Pricing." *Econometrica*, 80, 2469–2509.

Foucault, Thierry (1999). "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market." *Journal of Financial Markets*, 2, 99–134.

Foucault, Thierry, Ohad Kadan, and Eugene Kandel (2005). "The Limit Order Book as a Market for Liquidity." *Review of Financial Studies*, 18, 1171–1217.

Glosten, Lawrence R. (1994). "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49, 1127–1161.

Goettler, Ronald L., Christine A. Parlour, and Uday Rajan (2009). "Informed Traders and Limit Order Markets." *Journal of Financial Economics*, 93, 67–87.

Green, Richard, Burton Hollifield, and Norman Schurhoff (2007). "Financial Intermediation and the Costs of Trading in an Opaque Market." *Review of Financial Studies*, 20, 275–314.

Gromb, Dennis and Dimitri Vayanos (2002). "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs." *Journal of Financial Economics*, 66, 361–407.

Gromb, Dennis and Dimitri Vayanos (2010). "A Model of Financial Market Liquidity Based on Intermediary Capital." *Journal of the European Economic Association*, 8, 456–466.

Hansch, Oliver, Narayan Y. Naik, and S. Viswanathan (1998). "Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange." *Journal of Finance*, 53, 1623–1656.

Harris, Lawrence E. and Michael S. Piwowar (2006). "Secondary Trading Costs in the Municipal Bond Market Preview." *Journal of Finance*, 61, 1361–1397.

He, Zhiguo and Arvind Krishnamurthy (2012). "A Model of Capital and Crises." *Review of Economic Studies*, 79, 735–777.

Hendershott, Terrence and Albert Menkveld (2010). "Price Pressures." Forthcoming in *Journal of Financial Economics*. Available at http://ssrn.com/abstract=1411943.

Huang, Roger D. and Hans R. Stoll (1997). "The Components of the Bid-Ask Spread: A General Approach." *Review of Financial Studies*, 10, 995–1034.

Kaniel, Ron and Hong Liu (2006). "What Orders do Informed Traders Use?" *Journal of Business*, 79, 1867–1913.

Kumar, Praveen and Duane J. Seppi (1994). "*Limit and Market Orders with Optimising Traders*." Working paper, Carnegie Mellon University.

Lagos, Ricardo, Guillaume Rocheteau, and Pierre-Olivier Weill (2011). "Crises and Liquidity in Over-the-Counter Markets." *Journal of Economic Theory*, 146, 2169–2205.

Parlour, Christine A. and Duane J. Seppi (2008). "Limit Order Markets: A Survey." In *Handbook of Financial Intermediation and Banking*, edited by A. W. A. Boot and A. V. Thakor. Elsevier.

Ranaldo, Angelo (2004). "Order Aggressiveness in Limit Order Book Markets." *Journal of Financial Markets*, 7, 53–74.

Rinne, Kalle and Matti Suominen (2009). "A Structural Model of Short-Term Reversals." *SSRN*. Available at http://ssrn.com/abstract=1537922.

Rosu, Ioanid (2014). "Liquidity and Information in Order Driven Markets." Working paper, HEC, Paris. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1286193.

Reiss, Peter C. and Ingrid M. Werner (1998). "Does Risk Sharing Motivate Interdealer Trading?" *Journal of Finance*, 53, 1657–1703.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Online Appendices