Asset Heterogeneity, Market Fragmentation, and Quasi-Consolidated Trading*

Wei Li† Zhaogang Song‡

October 21, 2023

Abstract

Large asset heterogeneity is one of the most salient features of over-the-counter (OTC) markets. We first demonstrate, in a benchmark model, that asset heterogeneity results in market fragmentation, limiting the beneficial network externality of liquidity. We then introduce quasi-consolidated (QC) contract—a mechanism that enables assets of heterogeneous values to be traded at a uniform price—into the benchmark model and show that it increases total trading volume and social welfare by reducing market fragmentation. Nevertheless, the uniform pricing of QC trading leads to a cheapest-to-deliver effect, which harms liquidity for sellers who do not adopt QC trading; it also lowers profits for these sellers and even some sellers who adopt QC trading. Our model lays a foundation for analyzing liquidity and design in OTC markets of heterogeneous assets.

Keywords: Asset heterogeneity, Mortgage-backed Securities, Over-the-Counter Markets, Quasi-consolidated trading, TBA.

JEL Codes: G1, G11, G12, G21, D83, D53, D61

*We are grateful to Doug McManus at Freddie Mac and Marty Young at Goldman Sachs for helpful conversations about the MBS market. We appreciate the comments of Hank Bessembinder, Aytek Malkhozov, Vincent Maurin, Semih Üslü, and Shengxing Zhang, as well as seminar participants from Baruch College, Johns Hopkins Carey Business School, Virginal Tech, the 2021 MFA meeting, the 2023 Search and Matching Workshop in Asia-Pacific, and the 2023 EFA meeting.

†Zicklin School of Business at Baruch College. 137 East 22nd St, New York, NY 10010, USA. 1-646-660-6936. E-mail:wei.li@baruch.cuny.edu.

‡The Carey Business School of Johns Hopkins University. E-mail: zsong8@jhu.edu.

1 Introduction

Research on trading and liquidity in decentralized financial markets, also known as over-the-counter (OTC) markets, has progressed rapidly, especially since the 2008 financial crisis. The theoretical frameworks developed so far have analyzed various features of OTC markets, including searching, bargaining, dealer intermediation, network structure, and market opaqueness (see Duffie (2012) and Weill (2020) for recent surveys). In this paper, we study another salient feature of OTC markets—asset heterogeneity.

Substantial asset heterogeneity is prevalent in many important OTC markets like those of U.S. fixed-income securities. For example, in the corporate bond market, more than 100,000 bonds issued by about 5,600 firms are outstanding, as of 2022; in the municipal bond market, about 50,000 entities have issued more than 1.5 million bonds by the end of 2017 (Bessembinder, Spatt and Venkataraman, 2019); in the agency mortgage-backed securities (MBS) market, as of 2019, there are more than 800,000 securities backed by standard mortgage loans with different borrower, loan, and lender characteristics.

Such large asset heterogeneity is widely believed to hurt OTC market liquidity. Bessembinder et al. (2019) argue, for example, that "one reason that individual corporate bonds trade less frequently than equities is that an issuer often has multiple bond issues outstanding. While equity shares issued at different points in time by a given firm are fully substitutable, each bond issue is a separate contract with differing promised payments, maturity dates, and priority in case of default." Indeed, the average daily trading volume of corporate bonds and municipal bonds in 2019 was about \$34 and \$12 billions (bn), respectively, making an annual turnover in the range of 1.1-1.3. In contrast, the average daily trading volume of U.S. Treasury securities, which are much less heterogeneous (only around 400 securities are available), was about \$590 bn, making an annual turnover of about 12.8.

The agency MBS market, however, presents an intriguing exception: despite featuring highly heterogeneous assets, this market is actually "one of the most liquid fixed income markets in the world, with trading volumes typically in the trillions of dollars per year."[1] Specifically, its major trading mechanism, the to-be-announced (TBA) con-

[1]https://www.alliancebernstein.com/sites/library/Instrumentation/MORT-MAG-GR-EN-0118-FINAL.PDF

tract that trades a *cohort* of heterogeneous MBS at a uniform price, has an average daily volume of about $230 bn, making an annual turnover of about 10.7, comparable to that of U.S. Treasuries (Vickery and Wright, 2011; Gao, Schultz and Song, 2017). One might attribute the exceptional liquidity of TBA trading to the government backing of agency MBS rather than the TBA trading mechanism. Inconsistent with this conjecture, trading of agency MBS through standard individual-security based contracts—known as specified pools (SP) and similar to those used for corporate and municipal bonds—is inactive; it makes an annual turnover of 0.8, comparable to those of corporate and municipal bonds. Thus, it is mainly the TBA trading mechanism rather than the government backing that accounts for the remarkable liquidity of agency MBS.[2]

How does asset heterogeneity hurt OTC market liquidity? How does TBA trading improve liquidity in the presence of asset heterogeneity? Do these effects vary across assets and investors? Answering these questions is important not only for understanding agency MBS markets, which provide most funds for residential mortgages,[3] but also for potentially enhancing designs of other OTC markets. In fact, it has been conjectured that introducing a TBA-like trading mechanism can improve the liquidity of corporate bonds and municipal bonds.[4] In this paper, we build a theoretical framework to address these questions, which we hope can lay a foundation for analyzing liquidity and design in OTC markets of heterogeneous assets.

In the first part of our analysis, we develop a benchmark model to demonstrate how asset heterogeneity results in market fragmentation and hurts OTC market liquidity by restricting the positive network externality of liquidity.

Specifically, in the benchmark setup, assets of heterogeneous values are traded through standard contracts that fully specify the traded assets; we refer to such contracts as asset-specific (AS) contracts. Asset heterogeneity leads to market fragmentation on both the

[2]The patterns of liquidity differences across markets are similar if we use trading costs to measure liquidity; see Section 2 for details.

[3]As of the third quarter of 2020, 63% of the $11.5 trillion outstanding residential mortgage debt is securitized into agency MBS, based on the Urban Institute's reports.

[4]For example, Spatt (2004) mentions that "the analogy to the mortgage markets is instructive. Trading instruments based upon their main characteristics ... may be helpful and narrow the spreads." Moreover, Gao et al. (2017) argue that "corporate and municipal bonds trade in relatively illiquid over-the-counter markets. Parallel trading in the securities themselves and a forward contract on a generic security may increase the liquidity of those markets." Bessembinder et al. (2019) also ask whether there is "scope for the trading of packages of corporate bonds based on a set of prescribed characteristics."

selling side and the buying side because (1) a seller can sell only her own asset, and (2) although a buyer faces no constraints regarding which assets she can purchase,[5] she needs to pay a cost to participate in the trading of every asset. Such costs arise naturally when assets are heterogeneous: for example, before bidding for an asset, buyers need to collect data, predict future cash flows, and run models to analyze and value this asset. When this cost is higher, a buyer trades fewer assets; for simplicity, we consider the case in which the cost is so high that each buyer participates in the trading of at most one asset. In equilibrium, this AS-only market is fragmented into multiple segmented submarkets; on each submarket, an AS contract for a specific asset is used by a distinct subset of traders.

We model the trading process using a simple static search-and-matching framework. Specifically, on each trading venue (e.g., an AS submarket as defined above), buyers and sellers are randomly chosen and matched, and then a trade occurs between every matched buyer-seller pair. Accordingly, we define the *liquidity of an asset* as the probability of selling it successfully. We assume that the matching function features increasing returns to scale so that pooling more traders in the same trading venue enhances liquidity, which reflects the positive network externality on market liquidity often observed in OTC markets (Vayanos and Weill, 2008; Weill, 2020). With this feature, we show that asset heterogeneity results in market fragmentation, which limits the mass of traders on each trading venue, thereby harming liquidity.

In the second part of our analysis, we introduce a TBA-like contract to the benchmark setup and study how it affects liquidity in the presence of asset heterogeneity. Specifically, this TBA-like contract allows a seller to deliver *any* asset among a pre-defined cohort at the same price. That is, while an AS contract sets a price that is specific to the delivered asset, a TBA-like contract sets a uniform price for any deliverable asset within a cohort of heterogeneous assets. We call a TBA-like contract a quasi-consolidated (QC) contract because buyers pay the same price but may receive differing assets; in con-

[5] In practice, many investors have no preference for specific assets within a given asset class. For example, in the liquidity coverage ratio (LCR) requirements of Basel III, all MBS guaranteed by Fannie Mae and Freddie Mac are included in the level 2A category and receive the same haircut in computing the amount of "high-quality liquid assets" (Bank for International Settlements, 2013). Thus, banks have no preference for specific Fannie Mae and Freddie Mac MBS in meeting the LCR requirement. As also pointed out by Spatt (2004), "typically, buyers are not focused upon particular securities but instead are interested in purchasing a security with certain characteristics."

trast, security baskets and exchanged-traded funds are *fully*-consolidated because they combine a set of assets into one single security.

When both AS trading and QC trading are available, a seller who owns a QC-eligible asset (i.e., an asset that can be delivered to the QC contract), can sell her asset using either the AS contract specific for her asset or the QC contract. A buyer can choose the QC contract or any AS contract. As in the benchmark setup, we assume that a buyer needs to pay a cost to participate in the trading of one contract (the QC contract or an AS contract) and this cost is high such that each buyer participates in the trading of at most one contract. In equilibrium, the market is fragmented into one QC market and multiple AS submarkets, each of which consisting of a distinct subset of traders who use a distinct contract.

We focus on the parallel-trading equilibrium in which both AS and QC trading are used.[6] For QC trading, because all eligible assets are priced uniformly, low-value eligible assets are more likely to be delivered into the QC contract, which is often known as "cheapest-to-deliver" (CTD) practice. Because of this CTD effect, no seller would choose QC trading if it is (strictly) less liquid than AS trading; hence QC trading is (weakly) more liquid than AS trading in equilibrium. Thus, sellers who choose QC trading enjoy better liquidity but suffer price-discounts if their assets are more valuable than the uniform QC price, whereas sellers who choose AS trading experience worse liquidity but obtain individualized prices that are set specific for their assets. In equilibrium, sellers of QC-eligible assets choose QC trading only if their assets are less valuable than an endogenous threshold; if their assets more valuable than this threshold, they choose AS trading because, the price-discount of QC trading for their assets outweigh the liquidity benefit.

Comparing this parallel-trading market equilibrium with the AS-only market equilibrium in the benchmark setup, we show how introducing QC trading affects market liquidity, traders' profits, and social welfare.

In terms of market liquidity, we show that liquidity improves for assets that are traded on the (newly introduced) QC market but deteriorates for assets that remain on AS submarkets. First, for any given buyer-to-seller ratio, QC trading improves liquidity because of the network externality of concentrating more traders in one venue. Second, QC trad-

[6]There exist degenerate equilibria in which only AS or QC trading is used. These equilibrira are straightforward to analyze but are inconsistent with practice in markets like the agency MBS market.

ing attracts disproportionately more buyers than sellers to switch from AS trading to QC trading, which raises the buyer-to-seller ratio on the QC market and further increases the probability that assets on the QC market are sold successfully. Such disproportionate adoption of QC trading results from the difference in switching costs: intuitively, while switching is costless for buyers, it is costly for sellers who own high-value assets because of the CTD pricing. In addition, also because disproportionately more buyers than sellers leave AS trading, buyer-to-seller ratios decline on AS submarkets, hurting liquidity for assets that remain on AS submarkets.

Although introducing QC trading leads to varying liquidity effects across assets, it improves the *overall* market liquidity that we measure using the average trading probability of all assets. In our model, this probability is proportional to the total trading volume and turnover because the set of assets is fixed. Intuitively, because in the parallel-trading equilibrium multiple assets are traded together on the QC market, introducing QC trading partially "defragments" the benchmark AS-only market. Coupled with the network externality effect as captured by the increasing-return-to-scale matching function, this defragmentation reduces overall trading frictions and improves overall market liquidity.

In terms of traders' profits, we first show that after QC trading is introduced, all participating buyers are more likely to trade—those who switch from AS trading to QC trading benefit from returns to scale in liquidity and those who remain on AS submarkets benefit from lower buyer-to-seller ratios—and thus earn more expected profits. Moreover, buyers who choose not to participate in trading in the AS-only market (and earn zero expected profits) may choose to participate in trading in the parallel-trading market and earn positive expected profits. We then show that sellers who remain on AS submarkets earn less expected profits because they obtain the same prices but are less likely to trade. Finally, sellers who switch from AS trading to QC trading earn less (more) expected profits if their assets are more (less) valuable than an endogenous break-even level. Intuitively, although they switch to QC trading, they can be worse off because introducing QC trading makes AS submarkets less liquid and reduces the value of option to sell on AS submarkets.

Although introducing QC trading increases profits for some traders and decreases profits for others, we show that social welfare, which equals the total expected profits of all traders, is improved. On the one hand, the total profit of all buyers increases because every participating buyer earns weakly more expected profits and weakly more buyers participate. On the other hand, the total profit of all sellers also increases because (1) the total trading volume increases and (2) although sellers' profits differ, the *average* profit per trade across sellers remains unchanged.

Existing theoretical studies in the OTC market literature mostly model the trading of a single asset. In contrast, asset heterogeneity is the key feature of our model. Among the few existing works that consider markets of multiple assets, Vayanos and Wang (2007) and Vayanos and Weill (2008) also feature endogenous market fragmentation like ours.[7] However, they consider assets with the *same* value and explain why these assets can trade at different prices in particular. Our model focuses on assets with heterogeneous values and analyzes the effects of introducing a mechanism that can pool the trading of heterogeneous assets and that can be potentially used in various OTC markets.[8]

2 Background and Motivation

In this section, we briefly introduce the institutional background on OTC markets of U.S. fixed-income assets, focusing on the prevalence of asset heterogeneity and the potential of QC trading in improving market liquidity. We then discuss how the key economic mechanism in our model differs from existing (mostly informal) explanations for the liquidity of QC trading.

2.1 Institutional Background

U.S. fixed income markets provide a major source of financing for the economy. According to SIFMA, as of 2019, the outstanding balances for Treasury securities, agency MBS, corporate bonds, and municipal bonds were about \$16.7, \$7.7, \$9.6, and \$3.9 trillions

[7] Other studies include Weill (2008), Milbradt (2018), An (2019) and Üslü and Velioglu (2019).

[8] Our paper also adds to the existing and mostly empirical literature of agency MBS market structure and liquidity. See Bessembinder, Maxwell and Venkataraman (2013), Downing, Jaffee and Wallace (2009), Gao, Schultz and Song (2018), and Schultz and Song (2019) among others. To the best of our knowledge, we develop the first theoretical model that examines how TBA-like QC trading affects market liquidity for assets with heterogeneous values.

(tn), respectively (see the first column of Table 1).[9] Trading of these securities occur mostly in opaque, decentralized, and dealer-intermediated OTC markets rather than in centralized exchanges.[10]

A salient feature of these fixed-income markets is substantial asset heterogeneity. For example, in the corporate bond market, there are 105,132 bonds outstanding issued by 5607 firms, as of June 2022 (see the second column of Table 1). Moreover, according to Bessembinder et al. (2019), as of December 2017, there were over 1.5 million municipal bonds issued by about 50,000 issuers. It is widely believed that such substantial asset heterogeneity hurts market liquidity, as discussed in the Introduction. Consistent with this belief, the average daily trading volume of corporate bonds and municipal bonds is about \$34 and \$12 billions (bn), making an annual turnover of about 1.3 and 1.1, respectively. In contrast, for the U.S. Treasury market that features very low heterogeneity with only 404 securities (as of 2019), the average daily trading volume is about \$594 bn, making an annual turnover of about 12.8. Trading costs paint a similar picture: they are about 80-100 basis points (bps) for corporate bonds and municipal bonds (Di Maggio, Kermani and Song, 2017; Asquith, Covert and Pathak, 2019; Li and Schürhoff, 2019) but only about several bps for Treasury securities (Fleming et al., 2018; Song and Zhu, 2018).

The agency MBS market, however, presents an intriguing exception: despite featuring highly heterogeneous assets, markets for agency MBS are very active and liquid. In particular, while more than 800,000 agency MBS are outstanding (as of 2019), the average daily trading volume is about \$246 bn, resulting in an annual turnover of about 11.5 that is comparable to the annual turnover of Treasury securities (see the fourth row of Table 1).

Trading activities of agency MBS occur mainly through the TBA forward contracts. In particular, a TBA contract does not specify the specific MBS to be delivered at settlement; instead, it specifies, for example, that an MBS is deliverable if it is a Fannie Mae

[9] Other important but smaller fixed-income markets include those involving non-agency MBS (\$1.4 tn), federal agency securities (\$1.8 tn), and asset-backed-securities (\$1.8 tn).

[10] A small fraction of the fixed-income trading occurs in central limit-order books. In the U.S. Treasury market, for example, the inter-dealer segment of on-the-run securities trades through a centralized limit-order book run by BrokerTec (Fleming, Mizrach and Nguyen, 2018). Additionally, half of the inter-dealer trades of agency MBS are also executed on a centralized limit-order book run by TradeWeb (Schultz and Song, 2019).

Table 1. Summary of U.S. Fixed-Income Markets

Markets	Outstanding ($ tn)	Number of securities	Trading	
			Volume ($ bn)/Turnover	Cost (bp)
Municipal bond	3.9	1.5 million	12/1.1	80-100
Corporate bond	9.6	105,132	34/1.3	80-100
Treasury security	16.7	404	594/12.8	1-4
Agency MBS	7.7	824,462	246/11.5	
TBA			229/10.7	2-5
SP			17/0.8	20-60

This table provides aggregate summaries of U.S. fixed-income markets of municipal bonds, corporate bonds, Treasury securities, and agency MBS. The first column reports dollar outstanding amounts as of 2019 based on the SIFMA reports. The second column reports the number of securities outstanding for municipal bonds as of 2017 calculated by Bessembinder et al. (2019), for corporate bonds as of 2022 calculated using the Mergent Fixed Income Securities Database, for Treasury securities as of 2019 using the U.S. Treasury Monthly Statement of the Public Debt, and for agency MBS (standard ones backed by 15, 20, and 30 year fixed rate residential mortgage loans and guaranteed by Fannie Mae, Freddie Mac, and Ginnie Mae) as of 2019 calculated using the eMBS data. The third column reports the average daily trading volume in 2019 and the related annual turnover (equal to total trading volume in 2019 divided by outstanding amount as of 2019) based on the SIFMA reports. The fourth column reports the average trading costs estimated by Di Maggio et al. (2017) for corporate bonds, by Gao et al. (2017) for agency MBS, by Li and Schürhoff (2019) for municipal bonds, and by Fleming et al. (2018) for Treasury securities.

30-year fixed-rate MBS with a 4% security coupon rate.[11] It has been shown that TBA trading greatly improves trading activity and liquidity of agency MBS (Gao et al., 2017; Bessembinder et al., 2019). Indeed, the average daily TBA trading volume is $229 bn, which accounts for about 93% of the average daily trading volume of all agency MBS and makes an annual turnover of 10.7, while the TBA trading cost is about 2-5 bps. In contrast, the SP trading of agency MBS, which is the same as the standard trading of corporate bonds and municipal bonds on an individual-security basis, is quite illiquid. Its average daily trading volume is an order of magnitude lower than TBA trading and its trading costs are up to 60 bps (Gao et al., 2017)—both comparable to those of corpo-

[11]The specific MBS to deliver to an TBA contract is identified only two days before the settlement day. Details of the TBA settlement schedules and eligibility criteria are available at https://www.sifma.org/resources/general/mbs-notification-and-settlement-dates/ and https://www.sifma.org/wp-content/uploads/2017/06/uniform-practices-2019-chapter-8.pdf.

rate bonds and municipal bonds. This contrast implies that it is mainly the TBA trading mechanism (rather than the government backing of agency MBS) that accounts for the remarkable liquidity of the agency MBS market.

Because of its potential for improving market liquidity, TBA-like trading mechanism has been recommended for corporate and municipal bonds (Spatt, 2004; Gao et al., 2017; Bessembinder et al., 2019), as mentioned in Footnote 4. Moreover, given the importance of TBA trading for the agency MBS market, reforms have been implemented or proposed to further improve market liquidity. For example, in June 2019, the Federal Housing Finance Agency (FHFA) implemented the Single Security Initiative, under which Fannie Mae and Freddie Mac MBS are consolidated into "Uniform" MBS (UMBS). The single TBA contract for UMBS replaces the two TBA contracts for Fannie Mae and Freddie Mac MBS (Liu, Song and Vickery, 2020).

2.2 Discussion of the Economic Mechanism

The above institutional background and anecdotal evidence suggest that asset heterogeneity hurts OTC market liquidity and that the TBA-like cohort-based trading mechanism mitigates frictions associated with asset heterogeneity and improves liquidity. Before presenting our model of OTC market liquidity in the presence of asset heterogeneity, it is worth discussing and clarifying a few arguments that have been put forth *informally* on how TBA trading improves market liquidity (as mentioned in the Introduction, no theoretical models or any formal analyses have been done on this before).

First, prompted by the feature of the TBA contract that the specific MBS to be delivered is not disclosed to buyers at the time of the trade, several studies suggest that TBA trading improves liquidity by limiting information disclosure and reducing adverse selection.[12] Along this direction, Vickery and Wright (2011) argue that the TBA trading is analogous to the De Beers diamond auction and state that "the limits on information disclosure inherent in the TBA market seem to actually increase the market's liquidity by creating fungibility across securities and reducing information acquisition costs for buyers of MBS. A similar argument explains why De Beers diamond auctions involve selling pools of diamonds in unmarked bags that cannot be inspected by potential buyers."

[12] The idea that limiting information disclosure can reduce adverse selection and improve liquidity is reminiscent of the "Hirshleifer effect" (Hirshleifer, 1971).

Nevertheless, as French and McCormick (1984) pointed out, the De Beers selling strategy improves liquidity only if "potential buyers must believe that the asset is randomly selected from the prior distribution." This condition does not hold in TBA markets at all because sellers are well known to engage in "cheapest-to-deliver" practice; see Fusari, Li, Liu and Song (2022) for formal evidence.

Relatedly, Glaeser and Kallal (1997) build a formal model and show that liquidity of agency MBS can be improved when initial issuers limit information disclosure on MBS characteristics so that dealers cannot analyze such characteristics to gain substantial information advantage over investors. However, they model the liquidity of *all* agency MBS trading, which cannot explain why TBA trading is more liquid than SP trading. They do make an informal argument based on their model to explain the TBA market liquidity. Specifically, they argue that if an issuer "sells a mortgage bundle in that market, he gains more by withholding information than he would by explaining what exactly is in the bundle being sold. We think that this is one explanation why the TBA market exists." However, the information withheld by TBA contracts—which MBS will be delivered by issuers—will be revealed to dealers after the delivery; hence, TBA contracts *cannot* reduce the information advantage of dealers over investors.[13]

Second, in addition to using TBA to trade MBS, which we model, investors can also use TBA contracts for hedging because TBA contracts are forward contracts. For example, mortgage lenders often use TBA contracts to hedge interest rate risks (Vickery and Wright, 2011) and dealers use TBA contracts to hedge their MBS inventory (Chen, Liu, Sarkar and Song, 2023). Such hedging needs can further increase trading activities and improve the liquidity of TBA contracts. Nonetheless, while trades driven purely by hedging are usually settled by offsetting trades, a large proportion of TBA contracts are indeed settled by physical delivery of MBS, suggesting that TBA is an important mechanism for investors to buy and sell MBS. Our study focuses on this important trading function of TBA contracts.

[13]Glaeser and Kallal (1997) also raise another possible view that only "the 'worst' possible mortgage bundle" is traded in the TBA market, which they assume is of known value to all traders and hence perfectly liquid. This view is at odds with market practice that "a significant volume of physical delivery of securities occurs through the TBA market because, for many securities, the liquidity value of TBA trading generally exceeds any adverse selection discount implied by cheapest-to-deliver pricing" (Vickery and Wright, 2011). In fact, An, Li and Song (2022) show that a large number of highly heterogeneous MBS (e.g., over one-third of newly issued MBS) are sold through TBA contracts.

Third, we take asset heterogeneity as given and focus on the design of the trading mechanism. Security issuers can, of course, mitigate the frictions caused by asset heterogeneity by designing less heterogeneous securities (Gorton and Pennacchi, 1990; Subrahmanyam, 1991; DeMarzo, 2005). For example, Fannie Mae's Supers program allows investors to bundle various existing MBS into a single security. We differ by studying how a particular trading mechanism can help improve liquidity without changing the securities being issued.[14]

3 Asset Heterogeneity and Market Fragmentation

In this section, we develop a benchmark model with only standard asset-specific (AS) trading and show how asset heterogeneity results in market fragmentation and restricts market liquidity.

3.1 Setup

Assets are traded bilaterally between a continuum of risk-neutral buyers and a continuum of risk-neutral sellers.[15] We assume that the mass of buyers equals B and there are S assets, denoted by

$$\mathscr{A} := \{1, 2, \cdots, S\}. \tag{1}$$

We assume, for simplicity, that every asset has the same mass normalized to 1; thus, the total mass of assets equals S. Furthermore, we assume that every seller owns one share of a specific asset, so the mass of sellers who own each asset also equals 1 and the total mass of sellers equals S. In addition, we assume for simplicity that while a buyer can choose any asset, she can purchase up to one share of this asset. The mass of buyers who

[14]A related design of trading mechanism called portfolio trading has become popular in recent years (Li, O'Hara, Rapp and Zhou, 2023). In a portfolio trade, an investor sends a *single order* that consists of a basket of bonds to multiple dealers and execute the order with one dealer who can fill the entire order. One important difference is that a seller needs to have *all* bonds in the basket to use portfolio trading, while a seller who has *any* eligible MBS can use TBA trading.

[15]In Internet Appendix A.3, we allow buyers to be risk-averse and show that all our main results remain similar.

participate in trading can be less than B because buyers may choose not to participate.

Sellers and buyers value assets differently. Specifically, a share of asset j, for any $j \in \mathscr{A}$, is worth v_j to sellers and $v_j + \delta$ to buyers, where

$$v_j \overset{\text{iid}}{\sim} F \text{ with support } \mathcal{V} = [v_{\min}, v_{\max}]. \tag{2}$$

Therefore, a trade between a seller and a buyer results in a trading gain of $\delta > 0$.

Trades occur through standard AS contracts, with each AS contract allowing one asset to be delivered. For convenience of exposition, we say that traders are on the same trading venue (an AS submarket) if they trade with the same AS contract. Two types of trading frictions exist, both of which are related to asset heterogeneity.

The first type of friction, as in standard OTC models, is the search friction. We use, for simplicity, a static search-and-matching setup to capture search friction. Buyers and sellers on any trading venue are first randomly chosen-and-matched, and then trade; unmatched agents cannot trade. Formally, if sellers of mass m_s and buyers of mass m_b participate in a trading venue, the expected measure of buyer-seller matches on this venue equals

$$V(m_s, m_b) = \lambda \cdot (m_s m_b)^{\frac{1+\theta}{2}}, \tag{3}$$

where exogenous parameters λ measures the matching efficiency and $\theta \in (0,1)$ captures the liquidity benefit resulting from pooling multiple traders in one venue. We measure the liquidity level of a trading venue by the probability a seller on this venue is matched with a buyer, which, based on Eq. (3), equals

$$\pi^{\text{sell}} = \frac{V(m_s, m_b)}{m_s} = \lambda \frac{m_b^{\frac{1+\theta}{2}}}{m_s^{\frac{1-\theta}{2}}} = \lambda \left(\frac{m_b}{m_s}\right)^{\frac{1-\theta}{2}} m_b^{\theta}. \tag{4}$$

Hence, the liquidity level of a trading venue increases with the buyer-to-seller ratio (m_b/m_s) and the mass of buyers (m_b) on this venue. Also based on Eq. (3), the probability a buyer is matched with a seller equals

$$\pi^{\text{buy}} = \frac{V(m_s, m_b)}{m_b} = \lambda \frac{m_s^{\frac{1+\theta}{2}}}{m_b^{\frac{1-\theta}{2}}} = \lambda \left(\frac{m_s}{m_b}\right)^{\frac{1-\theta}{2}} m_s^{\theta}. \tag{5}$$

We assume that λ is so low that the matching probabilities π^{sell} and π^{buy} are less than 100%; otherwise, the market would be perfectly liquid.

Importantly, the assumption that $\theta > 0$ implies that liquidity improves when more traders choose the same trading venue. If, for example, the masses of sellers and buyers (m_s and m_b) on a venue increase proportionally, all traders on this trading venue experience better liquidity because both π^{sell} and π^{buy} increase according to Eqs. (4) and (5). In particular, if assets are homogeneous, then all sellers and all buyers would naturally trade together, which maximizes the liquidity externality of pooling traders in one venue; the probabilities that each seller and each buyer trade would equal, respectively,

$$\pi^s_{\text{ho}} = \lambda \left(\frac{B}{S}\right)^{\frac{1+\theta}{2}} S^{\theta} \text{ and } \pi^b_{\text{ho}} = \lambda \left(\frac{S}{B}\right)^{\frac{1-\theta}{2}} S^{\theta}. \tag{6}$$

We will show that when assets are heterogeneous, traders are naturally segmented into separate trading venues and the liquidity levels in Eq. (6) are generally not reached.

The second type of friction arises because it is costly for buyers to participate in the trading of an asset (Vayanos and Wang, 2013). When assets differ in value, such participation costs are natural because before a buyer bids, she needs to conduct costly valuation analysis that may involve collecting data, predicting future cash flows, and building pricing models (Eisfeldt, Lustig and Zhang, 2019). We denote by c the cost for a buyer to participate in every trading venue. Intuitively, a buyer would participate in fewer trading venues when c increases. For simplicity, we assume c is sufficiently high such that each buyer participates in the at most one trading venue (we will specify the required level of c in Assumption 1).

In addition, to determine the transaction price and allocation of trading gain, we assume that once a seller and a buyer are matched, nature chooses one side to make a take-it-or-leave-it trading proposal to the other side. The buyer is chosen with probability $\rho \in (0,1)$ and the seller is chosen with probability $1-\rho$. Thus, the transaction price of a share of type-j asset equals

$$P_{\text{as}}(v_j) = \begin{cases} v_j & \text{with probability } \rho, \\ v_j + \delta & \text{with probability } 1-\rho. \end{cases} \tag{7}$$

Thus, once a seller and a buyer are matched, the buyer expects to earn $\rho\delta$ and the seller expects to earn $(1-\rho)\delta$. The probability ρ thus reflects buyers' bargaining power against sellers.

3.2 Equilibrium

Next, we describe traders' choices and the equilibrium.

Sellers are naturally segmented into S different AS submarkets because a seller can use only the contract that allows her asset to be delivered. Therefore, the set of AS submarkets $\mathscr{M}_{\text{as}}$ is identical to the set of assets $\mathscr{A}$, i.e.,

$$\mathscr{M}_{\text{as}} = \mathscr{A} = \{1, 2, \cdots, S\}, \tag{8}$$

where the mass of sellers $s_j = 1$ on any AS submarket $j \in \mathscr{M}_{\text{as}}$. We denote the mass of buyers on AS submarket j by b_j and the mass of buyers who do not participate in any venue by b_0.

On each AS submarket, a seller would trade if she is matched to a buyer. Eq. (7) then implies that a seller on AS submarket j expects to earn

$$\psi_j^s = \pi_j^s \mathbf{E}\left[P_{\text{as}}(v_j) - v_j\right] = \pi_j^s (1-\rho)\delta, \tag{9}$$

where

$$\pi_j^s = \lambda \left(b_j\right)^{\frac{1+\theta}{2}} \tag{10}$$

equals the selling probability on this venue based on Eq. (4).

Each buyer maximizes her expected profit by choosing her trading venues in the set

$$\mathscr{C}^b := \{0\} \cup \mathscr{M}_{\text{as}}, \tag{11}$$

where 0 represents not participating in any trading venue. By choosing venue $j \in \mathscr{C}^b$, a buyer earns the expected profit

$$\psi_j^b = \pi_j^b \rho\delta - c \cdot \mathbb{1}_{\{j \neq 0\}}, \tag{12}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function and π_j^b equals the probability that this buyer on submarket j trades. As Eq. (5) implies, π_j^b equals

$$\pi_j^b = \begin{cases} \lambda\left(\frac{1}{b_j}\right)^{\frac{1-\theta}{2}} & \forall j \in \mathscr{M}_{\text{as}}, \\ 0 & j = 0. \end{cases} \tag{13}$$

Note that non-participating buyers (who choose $j = 0$) earn zero profits because they neither trade nor pay participation cost.

We assume that, if a buyer participates in multiple AS submarkets and is matched to sellers on more than one submarket, she randomly chooses one to trade. As mentioned above, we consider mainly the case in which each buyer participates in at most one AS submarket. For this, we impose the condition $c > 0.25\rho\delta$.[16] Moreover, we assume $c < \rho\delta$ to rule out a trivial no-trade equilibrium in which no one trades: if $c \geq \rho\delta$, a participating buyer's expected profit $\pi_j^b \rho\gamma - c$ would always be negative because the probability to buy $\pi_j^b < 1$. Together, we assume that the participation cost c satisfies

Assumption 1. A buyer's cost to participate in one venue $c \in (0.25\rho\delta, \rho\delta)$.

With these assumptions, we define the equilibrium as follows.

Definition 1. The AS-only market reaches an equilibrium if a vector of buyer masses

[16]We show in Lemma A.1 of the appendix that each buyer chooses to participate in at most one venue if the participation cost exceeds $\rho\delta/4$ on every venue; this holds even if the costs vary across venues.

$\{b_j : j \in \mathscr{C}^b\}$ satisfy $\sum_{j \in \mathscr{C}^b} b_j = B$ and

$$\psi_j^b \geq \max_{k \in \mathscr{C}^b, k \neq j} \quad \psi_k^b, \quad \forall j \in \mathscr{C}^b, \tag{14}$$

where ψ_j^b and ψ_k^b are defined in Eq. (12).

This is a competitive equilibrium in the sense that every buyer takes as given equilibrium buyer masses $\{b_j : j \in \mathscr{C}^b\}$. In equilibrium, every buyer weakly prefers her chosen venue to other choices in $\mathscr{C}^b$. The definition implies that, in equilibrium, a buyer's expected profit is non-negative because she could choose not to participate ($j = 0$) and earn zero profit.

The equilibrium is as follows.

Theorem 1 (**AS-only equilibrium**)**.** *In equilibrium, the mass of buyers on venue j equals*

$$b_j = \begin{cases} b^* & \text{if } j \in \{1, 2, \cdots, S\}, \\ B - Sb^* & \text{if } j = 0, \end{cases} \tag{15}$$

where

$$b^* := \min\left\{\frac{B}{S}, \bar{b}\right\} \text{ and } \bar{b} := \left(\frac{\lambda \rho \delta}{c}\right)^{\frac{2}{1-\theta}}, \tag{16}$$

In equilibrium, the AS-only market is fragmented into S submarkets of the same size, each attracting sellers of one distinct asset, whose mass equals 1, and a disjoint set of buyers, whose mass equals b^*. Intuitively, buyers spread evenly across AS submarkets because they are identical ex ante. In particular, because each buyer expect to earn $\rho\delta$ from any successful trade, she chooses the venue that maximizes the probability that she trades. As Eq. (13) shows, this probability decreases if a submarket attracts more buyers because buyers on the same submarket effectively compete for the opportunity to trade with a unit mass of sellers. Thus, in equilibrium every submarket must attract the same mass of buyers so that a buyer on any submarkets trades at the same probability and earns the same expected profit.

Moreover, note from Eq. (16) that the mass of buyers on any AS submarket cannot exceed $\bar{b}$ because otherwise buyers would earn negative expect profits. It follows that,

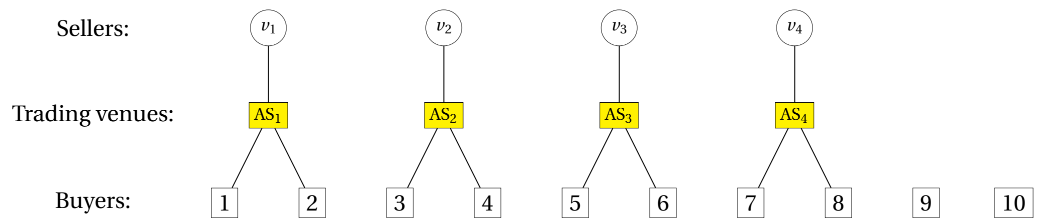


Figure 1. Illustration of Equilibrium in AS-only Market.

.

the total mass of buyers who participate in all AS submarkets

$$Sb^* \leq S\bar{b}. \tag{17}$$

In addition, based on Eqs. (15) and (16), the mass of non-participating buyers equals

$$b_0 = B - S \cdot \min\left\{\frac{B}{S}, \bar{b}\right\} = \max\left\{B - S\bar{b}, 0\right\}. \tag{18}$$

Thus, if the total mass of buyers B exceeds $S\bar{b}$, then buyers of mass $b_0 \geq 0$ would choose not to participate in trading and all buyers earn zero profit; if $B < S\bar{b}$, then all buyers would participate in the trading and earn positive profits.

We illustrate the equilibrium for a market with the total mass of assets $S = 4$ and total mass of buyers $B = 10$ in Fig. 1. In this case, the market fragments into 4 AS submarkets, each of which attracting buyers of mass $b^* = 2$. Mass of non-participating buyers $b_0 = 2$ because buyers 9 and 10 do not participate in any venue. Because in equilibrium each buyer earns the same profit, all buyers in this example earn zero profits.

3.3 Liquidity

As mentioned in Section 3.1, we measure the liquidity level of a venue by the probability that a seller on this venue can sell her asset successfully. Because every asset is traded on a distinct venue (AS submarket), this probability equals the probability that the asset on this venue is sold. Moreover, to capture the liquidity level of *the whole market*, we use the average probability that an asset is sold across all venues (AS submarkets). Because the total mass of assets S is fixed, this average probability is proportional to the total trading volume and turnover across all AS submarkets, both of which are commonly used to measure liquidity in practice. We derive the equilibrium liquidity levels as follows.

Corollary 1 (Liquidity level in AS-only equilibrium)**.** *In the equilibrium of the AS-only market, any asset j is sold with the same probability, i.e., $\pi_j^s = \pi^s$, where*

$$\pi^s = \lambda \left(b^*\right)^{\frac{1+\theta}{2}}. \tag{19}$$

Thus, the average probability that an asset is sold also equals

$$\pi^{\text{avg}} = \lambda \left(b^*\right)^{\frac{1+\theta}{2}}, \tag{20}$$

which implies that the total trading volume is

$$V = S \cdot \pi^{\text{avg}} = \lambda \left(b^*\right)^{\frac{1+\theta}{2}} \cdot S \tag{21}$$

Because every AS submarket consists of one unit mass of an asset and the same mass b^* of buyers, every AS submarket features the same level of liquidity π^s. Further, this implies, based on Eq. (9), that every seller earns the same expected profit, i.e., $\psi_j^s = \psi^s$ for any $j \in \mathscr{A}$, where

$$\psi^s = \pi^s(1-\rho)\delta. \tag{22}$$

Importantly, comparing the liquidity level in markets involving heterogeneous assets (Eq. (20)) with the liquidity level in markets involving homogeneous assets (Eq. (6)), we can understand how asset heterogeneity hurts liquidity. In particular, based on Eqs. (6), (16) and (20), we have

$$\pi^{\text{avg}} = \lambda \cdot \min\left\{\left(\frac{B}{S}\right)^{\frac{1+\theta}{2}}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{1+\theta}{1-\theta}}\right\} \le \lambda\left(\frac{B}{S}\right)^{\frac{1+\theta}{2}} \le \lambda\left(\frac{B}{S}\right)^{\frac{1+\theta}{2}} S^{\theta} = \pi_{\text{ho}}^s. \tag{23}$$

Thus, the loss in liquidity results from two effects. The first inequality reflects the effect of the cost c buyers need to pay to participate in the trading of an asset. When assets are heterogeneous, participation is costly for buyers because they need to analyze the assets they bid; when they expect low trading probability, they may choose not to participate, capping the mass of buyers on each submarket and weakly hurts liquidity. In contrast, when assets are homogeneous, buyers know the value of every asset and do not need to pay analysis costs ($c = 0$), so buyers would always participate. The second inequality in Eq. (23) reflects the liquidity externality of pooling traders. In particular, only 1 unit mass of an asset is traded on every AS submarket when assets are heterogeneous, whereas all S units of assets are traded on the same market when assets are homogeneous, i.e. $1 = 1^{\theta} \leq S^{\theta}$. Asset heterogeneity hence restricts the positive network externality on liquidity.

Moreover, Theorem 1 also implies that in equilibrium a buyer on any submarket trades with the same probability

$$\pi^b = \lambda \left(\frac{1}{b^*} \right)^{\frac{1-\theta}{2}}, \tag{24}$$

and hence earns the same expected profit

$$\psi^b = \pi^b \rho \delta - c. \tag{25}$$

We use these results in the next section.

4 Quasi-Consolidated Trading

In this section, we introduce QC trading to the benchmark model. We first describe the model setup for the parallel-trading market where both AS trading and QC trading are available, and then derive the equilibrium.

4.1 Setup

As discussed in Section 2.1, a TBA-like QC contract does not fully specify the deliverable asset but accepts a set of assets for delivery. Given our focus on asset heterogeneity in value, we assume that a QC contract specifies a threshold $\underline{v}$ such that assets are QC-eligible if and only if their value $v \geq \underline{v}$. To consider both QC-eligible and QC-ineligible assets, we assume $\underline{v} > v_{\min}$ so that not all assets are QC-eligible. In this section, we take $\underline{v}$ as given.

Once a QC contract is introduced, sellers of QC-eligible assets and all buyers can choose between the QC contract and AS contracts. Consequently, the set of assets $\mathscr{A}$ is endogenously partitioned into the set of assets traded on the QC market $\widetilde{\mathscr{M}}_{\text{qc}}$ and the set of assets traded on AS submarkets $\widetilde{\mathscr{M}}_{\text{as}}$.[17] That is,

$$\widetilde{\mathscr{M}}_{\text{as}} = \mathscr{A} - \widetilde{\mathscr{M}}_{\text{qc}}. \tag{26}$$

It follows that

$$S_{\text{as}} = S - S_{\text{qc}}, \tag{27}$$

where $S_{\text{qc}} = \left|\widetilde{\mathscr{M}}_{\text{as}}\right|$ and $S_{\text{as}} = \left|\widetilde{\mathscr{M}}_{\text{as}}\right|$ equal, respectively, the masses of assets on the QC market and on all AS submarkets. In addition, we denote the mass of buyers on the AS submarket for asset j by $\widetilde{b}_j$, the total mass of buyers on all AS submarkets by B_{as}, the mass of buyers on the QC market by B_{qc}, and the mass of buyers who do not participate in trading by $\widetilde{b}_0$. We maintain Assumption 1 so that every buyer participates in at most one venue; thus, $B_{\text{as}} = \sum_{j\in\widetilde{\mathscr{M}}_{\text{as}}} \widetilde{b}_j$ and $\widetilde{b}_0 = B - B_{\text{as}} - B_{\text{qc}}$.

Although in equilibrium buyers know the set of assets that will be delivered through QC contracts $\widetilde{\mathscr{M}}_{\text{qc}}$, a buyer on the QC market does not know which specific asset she will receive when she enters the QC contract. Because buyers are risk-neutral, we assume that the QC price equals

$$P_{\text{qc}} = \begin{cases} v_{\text{qc}} + \delta & \text{with probability } 1-\rho, \\ v_{\text{qc}} & \text{with probability } \rho, \end{cases} \tag{28}$$

where

$$v_{\text{qc}} = \mathbf{E}\left[v_\tau \middle| \tau \in \widetilde{\mathscr{M}}_{\text{qc}} \right] \tag{29}$$

equals the expected value of an asset traded on the QC market. A seller on the QC market will receive P_{qc} for delivering any QC-eligible asset. Eq. (28) implies that when a buyer

[17] We use tilde to indicate that a variable represents a characteristic of the market in which both AS and QC trading are available.

is matched, she expects to earn the same fraction ρ of the trading gain δ in a QC trade as in an AS trade; see Eq. (7) and discussions therein.

We assume that (both AS and QC) trading in the parallel-trading market features the same standard search frictions as in the AS-only market of Section 3, represented by the matching function specified in Eq. (3). Thus, the probabilities that a buyer and a seller on the QC market trade equal, respectively,

$$\widetilde{\pi}^b_{\text{qc}} = \lambda\left(\frac{S_{\text{qc}}}{B_{\text{qc}}}\right)^{\frac{1-\theta}{2}} S^\theta_{\text{qc}} \quad \text{and} \quad \widetilde{\pi}^s_{\text{qc}} = \lambda\left(\frac{B_{\text{qc}}}{S_{\text{qc}}}\right)^{\frac{1-\theta}{2}} B^\theta_{\text{qc}}; \tag{30}$$

the probabilities that a buyer and a seller trade on the AS submarket for asset j, equal, respectively,

$$\widetilde{\pi}^b_{\text{as},j} = \lambda\left(\frac{1}{\widetilde{b}_j}\right)^{\frac{1-\theta}{2}} \quad \text{and} \quad \widetilde{\pi}^s_{\text{as},j} = \lambda\left(\widetilde{b}_j\right)^{\frac{1+\theta}{2}}. \tag{31}$$

As before, non-participating buyers' probability of trading $\widetilde{\pi}^b_0 = 0$.

Moreover, we assume that for a buyer to participate in QC trading, she needs to pay c, the same cost she needs to pay to participate in the AS trading of an asset.[18] We maintain Assumption 1 to focus on the case in which each buyer participates in at most one venue—the QC market or one AS submarket.

4.2 Equilibrium

Next, we describe traders' choices and the equilibrium in the parallel-trading market.

As in the benchmark setup, a seller will trade if she is matched to a buyer. A seller of an QC-ineligible asset, whose value $v_j < \underline{v}$, has no choice but to sell on the AS submarket for asset j. Her expected profit is thus

$$\widetilde{\Psi}^s_{\text{as},j} = \widetilde{\pi}^s_{\text{as},j}\,\mathbf{E}\left[P_{\text{as}}(v_j) - v_j\right] = \widetilde{\pi}^s_{\text{as},j}(1-\rho)\delta. \tag{32}$$

[18]By assuming the same cost, we abstract away from the straightforward impact of participation cost difference on liquidity and instead, focus on the effects of network externality.

The choice set for a seller who owns a QC-eligible asset j is

$$\widetilde{\mathscr{C}}_j^s := \{\text{qc}, j\}, \tag{33}$$

where "qc" represents the QC market and "j" represents the AS submarket for asset j. Let $\varepsilon \in \widetilde{\mathscr{C}}_j^s$ represent the venue a seller chooses; then, this seller expects to earn

$$\widetilde{\psi}(\varepsilon, v_j) := \begin{cases} \widetilde{\psi}_{\text{as},j}^s = \widetilde{\pi}_{\text{as},j}^s (1-\rho)\delta & \text{if } \varepsilon = j, \\ \widetilde{\pi}_{\text{qc}}^s \mathbf{E}\left[P_{\text{qc}} - v_j\right] & \text{if } \varepsilon = \text{qc}, \end{cases} \tag{34}$$

where $\widetilde{\pi}_{\text{qc}}^s$ and $\widetilde{\pi}_{\text{as},j}^s$ are defined in Eqs. (30) and (31). A seller's profit on an AS submarket does not depend on the value of her asset v_j, whereas her profit on the QC market depends on the difference between the value of her asset v_j and the uniform price for all assets on the QC market P_{qc}.

We will show that in equilibrium the QC market is more liquid than any AS submarket. As a result, sellers whose assets are more valuable than P_{qc} face a trade-off between liquidity and price:while the QC market is more liquid, their assets could be sold at higher prices on AS submarkets. In equilibrium, sellers of high-value assets, whose values exceed the endogenous threshold $\bar{v}$, choose AS trading despite its illiquidity because their assets are much more valuable than P_{qc}.

Each buyer's choice set is

$$\widetilde{\mathscr{C}}^b := \{0, \text{qc}\} \cup \widetilde{\mathscr{M}}_{\text{as}}, \tag{35}$$

where 0 represents non-participation, "qc" represents the QC market, and $\widetilde{\mathscr{M}}_{\text{as}}$ represents the set of all AS submarkets (for both QC-eligible and QC-ineligible ones). A buyer who chooses k from $\widetilde{\mathscr{C}}^b$ expects to earn

$$\widetilde{\psi}_k^b = \begin{cases} \widetilde{\pi}_{\text{qc}}^b \mathbf{E}\left[v_\tau - P_{\text{qc}} \,\middle|\, \tau \in \widetilde{\mathscr{M}}_{\text{qc}}\right] - c & k = \text{qc} \\ \widetilde{\pi}_{\text{as},j}^b \mathbf{E}\left[v_j - P_{\text{as}}(v_j)\right] - c & k = j \in \widetilde{\mathscr{M}}_{\text{as}} \\ 0 & k = 0 \end{cases} \tag{36}$$

where $\widetilde{\pi}^b_{\text{qc}}$ and $\widetilde{\pi}^b_{\text{as},j}$ are defined in Eqs. (30) and (31). The pricing functions we assume in Eqs. (7) and (28) then imply that

$$\widetilde{\psi}^b_k = \widetilde{\pi}^b_k \rho\delta - c \cdot \mathbb{1}_{\{k \neq 0\}}, \tag{37}$$

where

$$\widetilde{\pi}^b_k = \begin{cases} \widetilde{\pi}^b_{\text{qc}} & k = \text{qc} \\ \widetilde{\pi}^b_{\text{as},j} & k = j \in \widetilde{\mathscr{M}}_{\text{as}} \\ 0 & k = 0. \end{cases} \tag{38}$$

Therefore, a buyer's expect profit depends her trading venue through trading probabilities.

To derive the equilibrium, we make a tie-breaking assumption that sellers prefer the QC market to an AS submarket when they are indifferent between the two. This assumption ensures that at least one asset is traded via the QC market ($S_{\text{qc}} \geq 1$).[19] We then state the equilibrium conditions.

Definition 2. The parallel-trading market reaches an equilibrium if sellers' venue choices $\{\varepsilon_j : j \in \mathscr{A}\}$ and buyer masses$\{b_k : k \in \mathscr{C}^b\}$ satisfy the following conditions:

- (Sellers of QC-ineligible assets) If a seller owns a QC-ineligible asset ($v_j < \underline{v}$), she trades on the AS submarket for her asset j.
- (Sellers of QC-eligible assets) If a seller owns a QC-eligible asset ($v_j \geq \underline{v}$), her trading venue choice ε_j maximizes her expected profit:

$$\widetilde{\psi}(\varepsilon_j, v_j) \geq \max_{e \in \widetilde{\mathscr{C}}^s_j, e \neq \varepsilon_j} \widetilde{\psi}(e, v_j) \tag{39}$$

  where $\widetilde{\psi}(\cdot, v_j)$ is defined in Eq. (34).

[19]The assumption also implies that all the units of an asset are traded either entirely through AS trading or entirely through QC trading.

- (Buyers) For any buyer who chooses $k \in \widetilde{\mathscr{C}}^b$, her expected profit

$$\widetilde{\psi}_k^b \geq \max_{k' \in \widetilde{\mathscr{C}}^b, k' \neq k} \widetilde{\psi}_{k'}^b, \tag{40}$$

  where $\widetilde{\psi}_k^b$ and $\widetilde{\psi}_{k'}^b$ are defined according to Eq. (37).

As in the benchmark model, this is a competitive equilibrium in the sense that every trader takes equilibrium buyer masses $\{b_k : k \in \mathscr{C}^b\}$ as given. The difference here is that each buyer has one more venue to choose—the QC market—in addition to AS submarkets and nonparticipating. Furthermore, trading probabilities of both buyers ($\widetilde{\pi}_k^b$) and sellers ($\widetilde{\pi}_{\text{qc}}^s$ and $\widetilde{\pi}_{\text{as},j}^s$) depend on the equilibrium set of sellers who choose the QC market and the equilibrium masses of buyers on each trading venue.

Multiple equilibria may arise because prices in the QC market can be self-fulfilling. In particular, if sellers expect buyers to bid high prices in the QC market, sellers of higher-value assets would choose the QC market, which can in turn justify buyers' high bids. We focus on the equilibrium with the highest QC price; in this equilibrium, the level of aggregate liquidity, measured by the total expected trading volume, reaches the maximum among all possible equilibria.

For ease of presenting the equilibrium, we define a function

$$\mu(x) := x\left(x^{\frac{2\theta}{1-\theta}} - 1\right), \tag{41}$$

which is non-negative and increasing when $x \geq 1$. The equilibrium is as follows.

Theorem 2 (**Parallel-trading equilibrium**). *Let*

$$\bar{v} = \max_{x \in [\underline{v}, v_{\max}]} \left\{ x : x \leq \mathbf{E}[v | v \in [\underline{v}, x]] + \left(1 - \frac{1}{\left(S \cdot \Pr\left\{v \in [\underline{v}, x]\right\}\right)^{\frac{2\theta}{1-\theta}}}\right)(1-\rho)\delta \right\} \tag{42}$$

and define

$$q := \Pr\left\{v \in [\underline{v}, \bar{v}]\right\} \tag{43}$$

$$\widetilde{b}^* := \min\left\{\frac{B}{S + \mu(Sq)}, \bar{b}\right\}, \tag{44}$$

where $\bar{b}$ is defined in Eq. (15). In the parallel-trading market, the equilibrium set of sellers' venue choices and the equilibrium vector of buyer masses are such that:

- *(Sellers) A seller chooses the QC market if her asset's value $v_j \in [\underline{v}, \bar{v}]$ and the AS submarket for asset j if $v_j \in [v_{\min}, \underline{v}) \cup (\bar{v}, v_{\max}]$, which implies that*

$$\widetilde{\mathscr{M}}_{\mathrm{qc}} = \left\{ j : j \in \mathscr{A} \text{ and } \underline{v} \le v_j \le \bar{v} \right\}, \tag{45}$$

$$\widetilde{\mathscr{M}}_{\mathrm{as}} = \mathscr{A} - \widetilde{\mathscr{M}}_{\mathrm{qc}}. \tag{46}$$

- *(Buyers) The mass of buyers who participate in the QC market equals*

$$B_{\mathrm{qc}} = (Sq)^{\frac{1+\theta}{1-\theta}} \cdot \widetilde{b}^{*}, \tag{47}$$

  the mass of buyers who participate in the AS submarket for asset j equals

$$\widetilde{b}_j = \widetilde{b}^{*} \quad \forall j \in \widetilde{\mathscr{M}}_{\mathrm{as}}, \tag{48}$$

  and the mass of buyers who do not participate in any trading venue equals

$$\widetilde{b}_0 = \max\left\{ B - \left(S + \mu(Sq) \right) \bar{b}, 0 \right\}. \tag{49}$$

A key endogenous parameter $\bar{v}$—the upper bound of the values of assets traded on the QC market—characterizes the equilibrium. A seller chooses the QC market only when the value of her asset falls in the interval $[\underline{v}, \bar{v}]$. Sellers of low-value assets, whose values are lower than $\underline{v}$, use AS trading because their assets are ineligible for QC trading; sellers of QC-eligible assets choose AS trading only if the values of their assets exceed $\bar{v}$. Given $\bar{v}$, we can find the fraction of assets sold on the QC market q. Then, the mass of sellers on the QC market, and the total mass of sellers on all AS submarkets equal, respectively,

$$S_{\mathrm{qc}} = Sq \quad \text{and} \quad S_{\mathrm{as}} = S(1-q). \tag{50}$$

In addition, because each AS submarket attracts buyers of mass $\widetilde{b}^{*}$, the total mass of

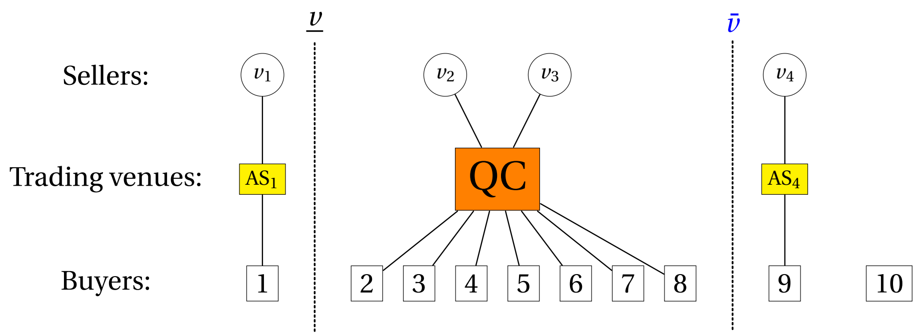


Figure 2. Illustration of the Parallel-Trading Equilibrium

buyers across all AS submarkets equals

$$B_{\text{as}} = S_{\text{as}} \widetilde{b}^*. \tag{51}$$

Fig. 2 illustrates an example parallel-trading equilibrium in which $q = 0.5$, $S = 4$, $B_{\text{qc}} = 7$, $\widetilde{b}^* = 1$, and $\widetilde{b}_0 = 1$. Because two assets, whose values equal v_2 and v_3, are traded together on the QC market, markets in this parallel-trading equilibrium are less fragmented than the AS-only equilibrium illustrated in Fig. 1.

Moreover, Fig. 2 also illustrates that the buyer-to-seller ratio on the QC market (which equals 3.5) exceeds the buyer-to-seller ratio on each AS submarket (which equals 1). In comparison, the buyer-to-seller ratio in the AS-only market as illustrated in Fig. 1 (which equals 2) lies between them. In addition, more buyers participate in trading in the parallel trading market than in the AS-only market ($9 > 9$). All these observations are true in general, as formally stated in the following result.

Corollary 2 (Comparison of market structure)**.** *As for buyer-to-seller ratios, we have that*

$$\widetilde{b}^* \le b^* \le B_{\text{qc}}/S_{\text{qc}}, \tag{52}$$

where $\widetilde{b}^*$ and $B_{\text{qc}}/S_{\text{qc}}$ equal, respectively, the ratio on each AS submarket and the ratio on the QC market in the parallel-trading market equilibrium, and b^* equals the ratio on each AS submarket in the AS-only equilibrium. Moreover, more buyers participate in *trading in the parallel-trading equilibrium than in the AS-only equilibrium, i.e.*

$$\widetilde{b}_0 \leq b_0. \tag{53}$$

To see why in the parallel-trading equilibrium the buyer-to-seller ratio on the QC market $B_{\text{qc}}/S_{\text{qc}}$ is weakly higher than the ratio on each AS submarket $\widetilde{b}^*$, consider the hypothetical situation in which $B_{\text{qc}}/S_{\text{qc}} < \widetilde{b}^*$. In this situation, the probability that a buyer on the QC market trades $\widetilde{\pi}^b_{\text{qc}}$ (given in Eq. (30)) would be strictly higher than the probability that a buyer on an AS submarket trades $\widetilde{\pi}^b_{\text{as},j}$ (given in Eq. (31)) because of both the strictly lower buyer-to-seller ratio ($B_{\text{qc}}/S_{\text{qc}} < \widetilde{b}^*$) and the liquidity externality of pooling multiple types of assets (i.e. $S^{\theta}_{\text{qc}} \geq 1$). In consequence, a buyer on the QC market would earn strictly higher profit than a buyer on an AS submarket, as seen from Eqs. (37) and (38), which cannot be an equilibrium because buyers on AS submarkets can switch costlessly to the QC market. Further, Corollary 2 also shows that introducing QC trading to the AS-only market lowers the buyer-to-seller ratio on AS submarkets but raises the ratio on the QC market, relative to the buyer-to-seller ratio on the AS-only market (b^*).

In addition, fewer buys choose not to participate ($\widetilde{b}_0 \leq b_0$), which can be derived by comparing Eqs. (18) and (49) given that $S + \mu(Sq) \geq S \geq 1$. Intuitively, more buyers to participate because each participating buyer earns weakly more profit: she is more likely to trade after paying the same participate cost.

4.3 Impact on Liquidity

In this section, we discuss the impact on liquidity of introducing QC trading. We first derive liquidity levels in the parallel-trading equilibrium and then compare them with the liquidity level in the AS-only equilibrium.

Liquidity levels in the parallel-trading equilibrium. As in the benchmark model, we measure the liquidity level of a trading venue by the probability that an asset on this venue is sold and measure the liquidity level of the whole market by the average probability that an asset is sold across all venues. Because the set of assets is fixed, this average probability is still proportional to the expected total trading volume and turnover across all venues.

We derive liquidity levels in the parallel-trading equilibrium based on Theorem 2 and report them as follows.

Corollary 3 (Liquidity levels in the parallel-trading equilibrium)**.** *In the parallel-trading equilibrium,*

- *an asset on any AS submarket (both QC-eligible and QC-ineligible) is sold with the same probability:* $\widetilde{\pi}^s_{\text{as},j} = \widetilde{\pi}^s_{\text{as}}$ *where*

$$\widetilde{\pi}^s_{\text{as}} = \lambda(\widetilde{b}^*)^{\frac{1+\theta}{2}}; \tag{54}$$

- *an asset on the QC market is sold with probability*

$$\widetilde{\pi}^s_{\text{qc}} = \lambda(\widetilde{b}^*)^{\frac{1+\theta}{2}}(Sq)^{\frac{2\theta}{1-\theta}}. \tag{55}$$

- *the average selling probability of all assets equals*

$$\widetilde{\pi}^{\text{avg}} = S(1-q)\widetilde{\pi}^s_{\text{as}} + Sq\widetilde{\pi}^s_{\text{qc}} = \lambda(\widetilde{b}^*)^{\frac{1+\theta}{2}}\frac{S+\mu(Sq)}{S}, \tag{56}$$

which implies that total expected trading volume across all venues equals

$$\widetilde{V} = S\widetilde{\pi}^{\text{avg}} = \lambda(\widetilde{b}^*)^{\frac{1+\theta}{2}}(S+\mu(Sq)). \tag{57}$$

We obtain Eq. (54) by plugging Eq. (48) into Eq. (31), and obtain Eq. (55) by plugging Eqs. (47) and (50) into Eq. (30). As Eq. (54) shows, in the parallel-trading equilibrium, AS submarkets (for both QC-eligible and QC-ineligible assets) are still equally liquid. Moreover, comparing Eq. (54) and Eq. (55), we find that $\widetilde{\pi}^s_{\text{as}} \le \widetilde{\pi}^s_{\text{qc}}$ (because $Sq \ge 1$); this is consistent with empirical findings regarding the agency MBS market (Gao et al., 2017).

Comparison of liquidity. Next, we examine the liquidity effects of introducing QC trading to the AS-only market by comparing the AS-only equilibrium presented in Theorem 1 and the parallel-trading equilibrium presented in Theorem 2.[20] We first present and explain the liquidity effects of introducing QC trading on individual assets.

Corollary 4 (**Comparison of liquidity of individual assets**). *Compared with the AS-only market, in the parallel-trading equilibrium, liquidity improves for assets that are traded on the QC market but worsens for assets that remain on AS submarkets. Formally,*

$$\widetilde{\pi}^s_{\text{as}} \le \pi^s \le \widetilde{\pi}^s_{\text{qc}}. \tag{58}$$

We derive the first inequality in Eq. (58) by comparing Eq. (54) with Eq. (19), knowing from Corollary 2 that $\widetilde{b}^* \le b^*$. That is, assets that remain on AS submarkets become less liquid because the buyer-to-seller ratio declines after QC trading is introduced. Thus, from the perspective of sellers who remain on AS submarket, the QC market siphons liquidity off AS submarkets.

We derive the second inequality in Eq. (58) as follows:

$$\widetilde{\pi}^s_{\text{qc}} = \lambda \left(\frac{B_{\text{qc}}}{S_{\text{qc}}}\right)^{\frac{1-\theta}{2}} \cdot B^{\theta}_{\text{qc}} \ge \lambda \left(b^*\right)^{\frac{1-\theta}{2}} \left(\frac{B_{\text{qc}}}{S_{\text{qc}}} \cdot S_{\text{qc}}\right)^{\theta} \ge \lambda \left(b^*\right)^{\frac{1-\theta}{2}} \left(b^* \cdot 1\right)^{\theta} = \pi^s. \tag{59}$$

In Eq. (59), the first equality is based on Eq. (30), the first inequality uses the increased buyer-to-seller ratio ($B_{\text{qc}}/S_{\text{qc}} \ge b^*$), the second inequality uses both the increased buyer-to-seller ratio ($B_{\text{qc}}/S_{\text{qc}} \ge b^*$) and the liquidity benefit of pooling more traders ($S^{\theta}_{\text{qc}} \ge 1$), and the last equality is based on Eq. (19).

We then examine the effects of introducing QC trading on overall liquidity of all assets. To do this, we introduce an auxiliary result concerning buyers. Specifically, plugging Eqs. (47), (48) and (50) into Eq. (31), we find that in equilibrium a buyer on any venue trades with the same probability, i.e., $\widetilde{\pi}^b_{\text{qc}} = \widetilde{\pi}^b_{\text{as},j} = \widetilde{\pi}^b$ for any $j \in \widetilde{\mathcal{M}}_{\text{as}}$, where

$$\widetilde{\pi}^b = \lambda \left(\frac{1}{\widetilde{b}^*}\right)^{\frac{1-\theta}{2}}. \tag{60}$$

[20] Note that when only one asset is traded on the QC market (i.e. $Sq = 1$), the parallel-trading described in Theorem 2 reduces to the AS-only equilibrium described in Theorem 1. The reason is that an QC market that attracts only one asset is equivalent to an AS submarket. Thus, we could also examine the effects of introducing QC trading by comparing the case in which $Sq \ge 1$ versus the cases in which $Sq = 1$ in Theorem 2.

Buyers are identical ex ante and pay the same participation cost c on any venue, so in equilibrium buyers on all venues must trade with the same probability and earn the same profit. It implies, based on Eqs. (24) and (60) and Corollary 2, that

$$\widetilde{\pi}^b = \lambda \left(\frac{1}{\widetilde{b}^*}\right)^{\frac{1-\theta}{2}} \geq \lambda \left(\frac{1}{b^*}\right)^{\frac{1-\theta}{2}} = \pi^b, \tag{61}$$

so a buyer on any venue is more likely to trade with the introduction of QC trading.

Lemma 1 (**Buyer participation and trading probability**). *After introducing QC trading, every participating buyer is more likely to trade ($\widetilde{\pi}^b \geq \pi^b$)*

With this auxiliary result, we present the overall liquidity effects on all assets of introducing QC trading.

Corollary 5 (**Comparison of liquidity of the whole market**). *After introducing QC trading, the overall liquidity of all assets improves. Formally,*

$$\widetilde{V} \geq V \textit{ and } \widetilde{\pi}^{\text{avg}} \geq \pi^{\text{avg}}. \tag{62}$$

The total expected trading volume equals the product of the probability that each participating buyer trades and the mass of all participating buyers. After QC trading is introduced, Corollary 2 implies that more buyers to participate in trading ($B - \widetilde{b}_0 \geq B - b_0$) and Lemma 1 shows that every participating buyer is more likely to trade ($\widetilde{\pi}^b \geq \pi^b$); thus, the total trading volume increases ($\widetilde{V} = (B - \widetilde{b}_0)\widetilde{\pi}^b \geq (B - b_0)\pi^b = V$). In addition, because the mass of assets S remains constant, the *average* probability that an asset is sold, which equals the total expected trading volume divided by S, also increases. That is, $\widetilde{\pi}^{\text{avg}} = \widetilde{V}/S \geq V/S = \pi^{\text{avg}}$.

In summary, introducing QC trading improves liquidity for assets that are now sold on the QC market but hurts liquidity for assets that remain on AS submarkets. Overall, the improvement for QC assets outweighs the harm to AS assets, so the average liquidity of all assets improves.

4.4 Impact on Trader Profits and Social Welfare

In this section, we examine the impact of introducing QC trading on traders' profits and social welfare.

First, we compare buyers' profits. Recall from Eq. (25) that a buyer on any venue expects to earn $\psi^b = \pi^b \rho\delta - c$ in the AS-only market. Moreover, Eqs. (37) and (60) imply that in the parallel-trading market equilibrium, a buyer on any venue earns the same expected profit, i.e., $\widetilde{\psi}^b_{\text{qc}} = \widetilde{\psi}^b_{\text{as},j} = \widetilde{\psi}^b$, where

$$\widetilde{\psi}^b = \widetilde{\pi}^b \rho\delta - c. \tag{63}$$

Lemma 1 shows that $\widetilde{\pi}^b \geq \pi^b$, so we have the following result.

Corollary 6 (Comparison of buyers' profits). *Compared with the AS-only market equilibrium, in the parallel-trading market equilibrium, a buyer earns more profits ($\widetilde{\psi}^b \geq \psi^b$).*

Second, we compare sellers' profits. Eq. (22) shows that in the AS-only market every seller earns $\psi^s = \pi^s(1-\rho)\delta$. Moreover, As Corollary 3 shows, in the parallel-trading equilibrium, a seller on any AS submarket, regardless of whether her asset is QC-ineligible or QC-eligible, trades with the same probability $\widetilde{\pi}^s_{\text{as}}$. Together with Eqs. (32) and (34), this implies that in the parallel-trading equilibrium, a seller on any AS submarket earns the same expected profit. That is, for any $j \in \widetilde{\mathcal{M}}_{\text{as}}$, $\widetilde{\psi}^s_{\text{as},j} = \widetilde{\psi}^s_{\text{as}}$, where

$$\widetilde{\psi}^s_{\text{as}} = \widetilde{\pi}^s_{\text{as}}(1-\rho)\delta. \tag{64}$$

Because $\widetilde{\pi}^s_{\text{as}} \leq \pi^s$ according to Corollary 4, we have $\widetilde{\psi}^s_{\text{as}} \leq \psi^s$, i.e., sellers who remain on AS markets earn less profits after QC trading is introduced because they are less likely to trade.

Further, Eqs. (28), (29) and (34) imply that in the parallel-trading equilibrium, a seller on the QC market expects to earn

$$\widetilde{\psi}^s(\text{qc}, v_j) = \widetilde{\pi}^s_{\text{qc}}\left(\mathbf{E}[P_{\text{qc}}] - v_j\right) = \widetilde{\pi}^s_{\text{qc}}\left((1-\rho)\delta + v_{\text{qc}} - v_j\right), \tag{65}$$

which decreases with the value of her asset v_j because any asset on the QC market receives the same price P_{qc}.

To compare the impact of introducing QC trading on sellers who switch to QC trading, we solve $\widetilde{\psi}^s(\text{qc}, v^*) = \psi^s$ and find that

$$v^* = v_{\text{qc}} + (1-\rho)\delta\left(1 - \frac{\pi^s}{\widetilde{\pi}^s_{\text{qc}}}\right). \tag{66}$$

That is, if a QC seller's asset is worth v^*, she earns the same profit as in the AS-only market. Because $\widetilde{\psi}^s(\text{qc}, v_j)$ decreases with v_j, a QC seller earns more expected profit than in the AS-only market ($\widetilde{\psi}^s(\text{qc}, v_j) > \psi^s$) if her asset is less valuable than v^* and earns less expected profit ($\widetilde{\psi}^s(\text{qc}, v_j) < \psi^s$) if her asset is more valuable that v^*. Intuitively, introducing QC trading reduces the value of the option to sell on the AS market. For sellers whose assets' values fall in the interval $(v^*, \bar{v}]$, such reduction outweighs the benefits of improved liquidity, so they are worse off despite their switching to QC trading.

Would the introduction of QC trading *always* hurt some QC sellers? It turns out, the answer is no. Specifically, it is possible that all QC sellers earn more profits than what they could have earned in the AS-only equilibrium (i.e. $v^* \geq \bar{v}$) if all QC eligible assets, including the most valuable ones, are sold through QC trading (i.e. $\bar{v} = v_{\max}$). Nonetheless, if the most valuable assets are sold through AS trading ($\bar{v} < v_{\max}$), as is the case in the agency MBS market (An et al., 2022), then $v^* \leq \bar{v}$ and QC sellers whose asset values fall in the interval $[v^*, \bar{v}]$ earn weakly less profits than what they could have earned in the AS-only market.

We summarize these results concerning sellers' profits formally in Corollary 7 and illustrate them in Fig. 3.

Corollary 7 (**Comparison of seller's profits**)**.** *Compared with the AS-only market equilibrium, in the parallel-trading market equilibrium, a seller on AS markets earns less profit ($\widetilde{\psi}^s_{\text{as}} \leq \psi^s_{\text{as}}$), a seller on the QC market earns more profit if her asset's value $v_j < v^*$ and earn less profit if $v_j \in (v^*, \bar{v}]$. If $\bar{v} < v_{\max}$, then $v^* \leq \bar{v}$.*

Finally, we compare social welfare, which aggregates all traders' expected profits. The social welfare in the AS-only market equilibrium equals

$$\Omega = \underbrace{(B - b_0)(\rho\delta\pi^b - c)}_{\text{all buyers' total profit}} + \underbrace{S\pi^{\text{avg}}(1-\rho)\delta}_{\text{all sellers' total profit}}, \tag{67}$$

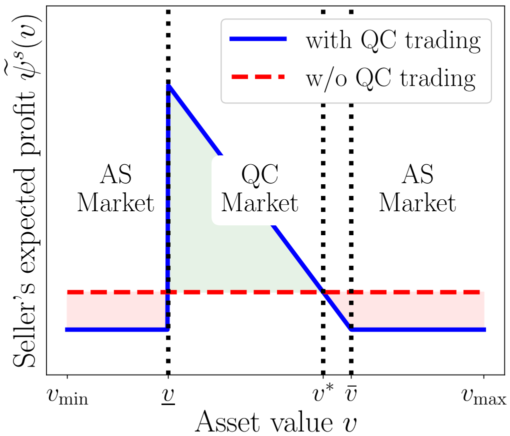


Figure 3. Impact of introducing QC trading on individual seller's profit when $v^* < \bar{v} < v_{\max}$.

where $B - b_0$ equals the total mass of participating buyers, $\rho\delta\pi^b - c$ equals the profit of each participating buyer (Eq. (25)), and $\pi^{\text{avg}}(1-\rho)\delta$ equals the expected profit of each seller (based on Eqs. (20) and (22)). Moreover, social welfare in the parallel-trading equilibrium equals

$$\begin{aligned}
\widetilde{\Omega} &= \underbrace{(B-\widetilde{b}_0)(\rho\delta\widetilde{\pi}^b - c)}_{\text{all buyers' total profit}} + \underbrace{S(1-q)\widetilde{\pi}^s_{\text{as}}(1-\rho)\delta}_{\text{AS sellers' total profit}} + \underbrace{Sq\widetilde{\pi}^s_{\text{qc}}\,\mathbf{E}\left[(1-\rho)\delta + v_{\text{qc}} - v_j | v_j \in [\underline{v}, \bar{v}]\right]}_{\text{QC sellers' total profit}} \\
&= \underbrace{(B-\widetilde{b}_0)(\rho\delta\widetilde{\pi}^b - c)}_{\text{all buyers' total profit}} + \underbrace{S(1-q)\widetilde{\pi}^s_{\text{as}}(1-\rho)\delta}_{\text{AS sellers' total profit}} + \underbrace{Sq\widetilde{\pi}^s_{\text{qc}}(1-\rho)\delta}_{\text{QC sellers' total profit}} \\
&= \underbrace{(B-\widetilde{b}_0)(\rho\delta\widetilde{\pi}^b - c)}_{\text{all buyers' total profit}} + \underbrace{S\widetilde{\pi}^{\text{avg}}(1-\rho)\delta}_{\text{all sellers' total profit}}
\end{aligned} \tag{68}$$

where $B - \widetilde{b}_0$ equals the total mass of participating buyers, $\rho\delta\widetilde{\pi}^b - c$ equals the profit of each participating buyer (Eq. (63)), $S(1-q)$ equals the total mass of sellers on AS submarkets, $\widetilde{\pi}^s_{\text{as}}(1-\rho)\delta$ equals the expected profit of each seller on AS submarkets (as Eq. (64) shows), Sq equals the total mass of sellers on the QC submarket, $\widetilde{\pi}^s_{\text{qc}}\,\mathbf{E}[(1-\rho)\delta + v_{\text{qc}} - v_j | v_j \in [\underline{v}, \bar{v}]]$ equals the average expected profit of QC sellers (note that the expectation is taken for $v_j \in [\underline{v}, \bar{v}]$), and $\widetilde{\pi}^{\text{avg}}$ equals the average asset selling probability based on Eq. (56). Because assets on the QC market are priced uniformly according to their average value v_{qc}, we have that $\mathbf{E}\left[(1-\rho)\delta + v_{\text{qc}} - v_j | v_j \in [\underline{v}, \bar{v}]\right] = (1-\rho)\delta$, so the average expected profit of QC sellers equals $\widetilde{\pi}^s_{\text{qc}}(1-\rho)\delta$.

Comparing Eqs. (67) and (68), we see that $\widetilde{\Omega} \geq \Omega$ because (1) as Lemma 1 shows, more buyers participate ($B - \widetilde{b}_0 \geq B - b_0$) and every participating buyer is more likely to trade ($\widetilde{\pi}^b \geq \pi^b$), and (2) as Corollary 5 shows, the overall market liquidity improves ($\widetilde{\pi}^{\text{avg}} \geq \pi^{\text{avg}}$). Thus, we have the following result:

Corollary 8 (**Comparison of Social Welfare**)**.** *Introducing QC trading improves social welfare ($\widetilde{\Omega} \geq \Omega$).*

5 Conclusion

We build a model of OTC markets involving assets with heterogeneous fundamental values. We first show that asset heterogeneity hurts liquidity by restricting the network externality of pooling traders together. Then, motivated by the remarkably liquid TBA trading of agency MBS, we analyze the TBA-like quasi-consolidated trading mechanism—under which heterogeneous assets within a cohort are sold at a uniform price. We find that introducing QC trading increases total trading volume and social welfare but hurts liquidity for assets that remain to be traded via standard asset-specific contracts.

Substantial asset heterogeneity is a salient and prevalent feature of OTC markets. Our theoretical framework lays a foundation for analyzing liquidity and trading design in various OTC markets involving heterogeneous assets, including corporate bonds, municipal bonds, MBS, and asset-backed securities. For example, our framework can be used to investigate a host of design issues, including whether to introduce QC trading and how to set the specification of QC trading.[21] Analyses along this direction would constitute important future developments.

[21] As a demonstration, in the Internet Appendix A.1, we conduct a simple analysis of the choice of QC-eligibility requirements, motivated by the restrictions imposed by the Securities Industry and Financial Markets Association (SIFMA) on TBA-eligibility of high-balance loans in 2008 (Vickery and Wright, 2011).

A Appendix: Proofs

Lemma A.1 (**Conditions for maximal buyer fragmentation**). *Let c_j denote a buyer's cost to participate in trading venue j in a set $\mathscr{M}$. If $c_j > \rho\delta/4$ for every j, a buyer participates in at most one venue.*

Proof of Lemma A.1. Without loss of generality, suppose that a buyer participates in multiple venues in the set $\mathscr{I} = \{1, \cdots, I\}$ and $\pi_1^b \leq \pi_i^b$ for $i \in \mathscr{I}$. The buyer expects to earn

$$\psi = \rho\delta\left(1 - (1 - \pi_1^b)\prod_{i=2}^{I}(1 - \pi_i^b)\right) - c_1 - \sum_{i=2}^{I} c_i. \tag{A.1}$$

If the buyer quits trading venue 1, she would earn

$$\psi' = \rho\delta\left(1 - \prod_{i=2}^{I}(1 - \pi_i^b)\right) - \sum_{i=2}^{I} c_i. \tag{A.2}$$

Because $\pi_1^b \leq \pi_i^b$, we have that

$$\frac{\psi' - \psi}{\rho\delta} = \frac{c_1}{\rho\delta} - \pi_1^b(1 - \pi_2^b)\prod_{i>2, i\in\mathscr{I}}(1 - \pi_i^b) \geq \frac{c_1}{\rho\delta} - \pi_2^b(1 - \pi_2^b) \geq \frac{c_1}{\rho\delta} - \frac{1}{4} > 0. \tag{A.3}$$

If a buyer participates in more than one trading venue, she could earn strictly more profit by quitting the venue with the lowest matching probability, so in equilibrium a buyer participates in at most one venue. □

Lemma A.2. *The mass of buyers on any AS submarket $b_j \leq \bar{b} = \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}$.*

Proof of Lemma A.2. If $b_j > \bar{b}$, then a buyer's profit $\pi_j^b\rho\delta - c = \lambda\rho\delta/b_j^{\frac{1-\theta}{2}} - c < 0$, which cannot be an equilibrium. □

Proof of Theorem 1. Assumption 1 implies, according to Lemma A.1, that a buyer participates in at most one submarket. If $b_j > 0$ and $b_k > 0$ for two submarkets j and k, then Eqs. (12) and (14) imply that $\pi_j^b\rho\delta - c = \pi_k^b\rho\delta - c$, which implies that $\pi_j^b = \pi_k^b$. Because $s_j = s_k = 1$, we have that $b_j = b_k$.

Lemma A.2 implies that the total mass of buyers participating in all S submarkets cannot exceed $S \cdot \bar{b}$. First, if $B \geq S\bar{b}$, then some buyers do not participate in any submarket and earn zero profit. Thus, participating buyers also earn zero profits, which implies that $\pi_j^b = c/(\rho\delta)$ and $b_j = \bar{b}$ for all $\forall j \in \mathscr{A}$. Second, if $B < S\bar{b}$, then every buyer earns the same and positive profit, which implies that $b_j = B/S \; \forall j$. Therefore, $b_j = b^* = \min\{B/S, \bar{b}\}$ for any submarket j, which implies that $b_0 = B - Sb^*$. □

Proof of Corollary 1. Follows directly from Eqs. (4) and (5) and Theorem 1. □

Lemma A.3. *If a buyer on any trading venue trades at the same probability, then* $\widetilde{b}^* = B_{\text{as}}/S_{\text{as}} = B_{\text{qc}}/S_{\text{qc}}^{\frac{1+\theta}{1-\theta}}$ *and* $\widetilde{\pi}_{\text{qc}}^s = \widetilde{\pi}_{\text{as}}^s S_{\text{qc}}^{\frac{2\theta}{1-\theta}}$.

Proof of Lemma A.3. If every buyer on AS submarkets experiences the same liquidity, then $\widetilde{b}_j = \widetilde{b}^* = B_{\text{as}}/S_{\text{as}}$ for any $j \in \mathscr{M}_{\text{as}}$. It follows that $\widetilde{\pi}_{\text{as}}^b = \lambda\left(1/\widetilde{b}^*\right)^{\frac{1-\theta}{2}}$ and $\widetilde{\pi}_{\text{as}}^s = \lambda\left(\widetilde{b}^*\right)^{\frac{1+\theta}{2}}$. Together with Eq. (30), we have that

$$\frac{\widetilde{\pi}_{\text{as}}^b}{\widetilde{\pi}_{\text{qc}}^b} = \frac{S_{\text{as}}^{\frac{1-\theta}{2}} B_{\text{qc}}^{\frac{1-\theta}{2}}}{B_{\text{as}}^{\frac{1-\theta}{2}} S_{\text{qc}}^{\frac{1+\theta}{2}}} = \left(\frac{S_{\text{as}} B_{\text{qc}}}{B_{\text{as}} S_{\text{qc}}^{\frac{1+\theta}{1-\theta}}}\right)^{\frac{1-\theta}{2}} \tag{A.4}$$

and

$$\frac{\widetilde{\pi}_{\text{as}}^s}{\widetilde{\pi}_{\text{qc}}^s} = \frac{B_{\text{as}}^{\frac{1+\theta}{2}}}{S_{\text{as}}^{\frac{1+\theta}{2}}} \frac{S_{\text{qc}}^{\frac{1-\theta}{2}}}{B_{\text{qc}}^{\frac{1+\theta}{2}}} = \left(\frac{\widetilde{\pi}_{\text{qc}}^b}{\widetilde{\pi}_{\text{as}}^b}\right)^{\frac{1+\theta}{1-\theta}} \frac{1}{S_{\text{qc}}^{\frac{2\theta}{1-\theta}}}. \tag{A.5}$$

If $\widetilde{\pi}_{\text{as}}^b = \widetilde{\pi}_{\text{qc}}^b$, then $\widetilde{b}^* = B_{\text{as}}/S_{\text{as}} = B_{\text{qc}}/S_{\text{qc}}^{\frac{1+\theta}{1-\theta}}$ and $\widetilde{\pi}_{\text{qc}}^s = \widetilde{\pi}_{\text{as}}^s S_{\text{qc}}^{\frac{2\theta}{1-\theta}}$. □

Lemma A.4. *A seller of asset j earns more profit through QC trading than through AS trading if and only if the value of her asset* $v_j \leq v_{\text{qc}} + \left(1 - \frac{\widetilde{\pi}_{\text{as},j}^s}{\widetilde{\pi}_{\text{qc}}^s}\right)(1-\rho)\delta$.

Proof of Lemma A.4. Eq. (34) implies that $\widetilde{\psi}_{\text{as},j}^s = \widetilde{\pi}_{\text{as},j}^s(1-\rho)\delta$ and a QC seller earns $\widetilde{\psi}_{\text{qc}}^s = \widetilde{\pi}_{\text{qc}}^s\left(\mathbf{E}[P_{\text{qc}}] - v_j\right) = \widetilde{\pi}_{\text{qc}}^s\left(v_{\text{qc}} + (1-\rho)\delta - v_j\right)$. Thus, their difference equals

$$\widetilde{\psi}_{\text{qc}}^s - \widetilde{\psi}_{\text{as},j}^s = \widetilde{\pi}_{\text{qc}}^s(v_{\text{qc}} - v_j) + (\widetilde{\pi}_{\text{qc}}^s - \widetilde{\pi}_{\text{as},j}^s)(1-\rho)\delta, \tag{A.6}$$

which is positive only if and only if $v_j \leq v_{\text{qc}} + \left(1 - \frac{\widetilde{\pi}^s_{\text{as},j}}{\widetilde{\pi}^s_{\text{qc}}}\right)(1-\rho)\delta$. □

Proof of Theorem 2. First, we derive buyer masses. In equilibrium buyers earn the same expected profit, which implies that all participating buyers trade at the same probability. It implies based on Lemma A.3, that the mass of buyers on every AS submarket equals $\widetilde{b} = B_{\text{as}}/S_{\text{as}} = B_{\text{qc}}/S_{\text{qc}}^{\frac{1+\theta}{1-\theta}}$. Thus, the total mass of participating buyers

$$B - \widetilde{b}_0 = B_{\text{as}} + B_{\text{qc}} = \widetilde{b}\left(S_{\text{as}} + S_{\text{qc}}^{\frac{1+\theta}{1-\theta}}\right) = \widetilde{b}\big(S + \mu(Sq)\big) \leq B, \tag{A.7}$$

which implies that $\widetilde{b} \leq \frac{B}{S+\mu(Sq)}$. In addition, Lemma A.2 shows that $\widetilde{b} \leq \bar{b}$. Thus,

$$\widetilde{b} \leq \widetilde{b}^* := \min\left\{\frac{B}{S+\mu(Sq)}, \bar{b}\right\}. \tag{A.8}$$

If $\widetilde{b} < \widetilde{b}^*$, then some buyers do not participate ($\widetilde{b}_0 > 0$) and yet participating buyers earn strictly positive profits ($\widetilde{\pi}^b \rho\delta - c > 0$), which cannot be an equilibrium. Thus, $\widetilde{b} = \widetilde{b}^*$.

Second, we find sellers' venue choices. Define

$$\eta(v^*) := \mathbf{E}[v|v \in [\underline{v}, v^*]] + (1-\rho)\delta\left(1 - \frac{1}{\big(S\cdot \Pr\{v \in [\underline{v}, v^*]\}\big)^{\frac{2\theta}{1-\theta}}}\right), \tag{A.9}$$

which is an increasing function. If an asset's value $v' > \bar{v}$, then by definition of $\bar{v}$ in Eq. (42), $v' > \eta(v') \geq \eta(\bar{v})$; if an asset's value $v' \leq \bar{v}$, then $v' \leq \bar{v} \leq \eta(\bar{v})$. Lemmas A.3 and A.4 imply that a seller prefers the AS market if

$$v > v_{\text{qc}} + \left(1 - \frac{\widetilde{\pi}^s_{\text{as}}}{\widetilde{\pi}^s_{\text{qc}}}\right)(1-\rho)\delta = v_{\text{qc}} + (1-\rho)\delta\left(1 - \frac{1}{S_{\text{qc}}^{\frac{2\theta}{1-\theta}}}\right). \tag{A.10}$$

Thus, a seller prefers the AS market if $v > \bar{v}$ and the QC market otherwise. Sellers of QC-ineligible assets can use only AS trading. Therefore, sellers choose the QC market if $v \in [\underline{v}, \bar{v}]$ and the AS market otherwise. □

Lemma A.5. $1 \leq \frac{S+\mu(Sq)}{S} \leq (Sq)^{\frac{2\theta}{1-\theta}}$ *and inequalities bind when* $Sq = 1$.

Proof of Lemma A.5. By definition, $q \leq 1$, $\mu(1) = 0$, and $Sq \geq 1$. It follows that $\mu(Sq) =$

$Sq\left((Sq)^{\frac{2\theta}{1-\theta}}-1\right) \le S\left((Sq)^{\frac{2\theta}{1-\theta}}-1\right)$, which implies that $(Sq)^{\frac{2\theta}{1-\theta}} \ge 1+\frac{\mu(Sq)}{S} = \frac{S+\mu(Sq)}{S} \ge 1$. When $Sq=1$, $\mu(Sq)=0$ and $S+\mu(Sq)=S$. □

Proof of Corollary 3. Described in the main text right after Corollary 3. □

Proof of Corollaries 2, 4 and 5 and Lemma 1. Eqs. (16) and (44) imply that $\widetilde{b}^* \le b^*$. Eq. (47) implies that $B_{\text{qc}}/S_{\text{qc}} = \widetilde{b}^* S_{\text{qc}}^{\frac{2\theta}{1-\theta}} \ge \widetilde{b}^*$. Eqs. (18) and (49) imply that $b_0 \ge \widetilde{b}_0$; Eqs. (19) and (54) imply that $\widetilde{\pi}^s_{\text{as}} \le \pi^s$; Eqs. (24) and (60) imply that $\pi^b \le \widetilde{\pi}^b$. Thus, $\widetilde{V} = (B-\widetilde{b}_0)\widetilde{\pi}^b \ge (B-b_0)\pi^b = V$ and $\widetilde{\pi}^{\text{avg}} = \widetilde{V}/S \ge V/S = \pi^{\text{avg}}$. In addition, Lemma A.5 implies that $S\cdot(Sq)^{\frac{2\theta}{1-\theta}} \ge S+\mu(Sq)$, which implies that

$$\frac{B}{S+\mu(Sq)}(Sq)^{\frac{2\theta}{1-\theta}} = \frac{B}{S}\frac{S}{S+\mu(Sq)}(Sq)^{\frac{2\theta}{1-\theta}} \ge \frac{B}{S}. \tag{A.11}$$

Thus,

$$\begin{aligned}\frac{B_{\text{qc}}}{S_{\text{qc}}} &= \widetilde{b}^* S_{\text{qc}}^{\frac{2\theta}{1-\theta}} = \min\left\{\frac{B}{S+\mu(Sq)}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}\right\}(Sq)^{\frac{2\theta}{1-\theta}} \\ &\ge \min\left\{\frac{B}{S+\mu(Sq)}(Sq)^{\frac{2\theta}{1-\theta}}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}\right\} \ge \min\left\{\frac{B}{S}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}\right\} = b^*\end{aligned} \tag{A.12}$$

Thus, $\widetilde{b}^* \le b^* \le B_{\text{qc}}/S_{\text{qc}}$. In addition, Eq. (59) shows that $\widetilde{\pi}^s_{\text{qc}} \ge \pi^s$. □

Proof of Corollaries 6 to 8. We show by contradiction that $v^* \in (\bar{v}, v_{\max}]$ cannot hold. If $v^* \in (\bar{v}, v_{\max}]$, then Eq. (42) implies that

$$v^* = v_{\text{qc}} + (1-\rho)\delta\left(1-\frac{\pi^s}{\widetilde{\pi}^s_{\text{qc}}}\right) > v_{\text{qc}} + (1-\rho)\delta\left(1-\frac{\widetilde{\pi}^s_{\text{as}}}{\widetilde{\pi}^s_{\text{qc}}}\right), \tag{A.13}$$

which implies that $\pi^s_{\text{as}} > \pi^s$; it contradicts our result that $\pi^s_{\text{as}} > \pi^s$. Thus, if $\bar{v} < v_{\max}$, then $v^* \le \bar{v}$. Proofs of other results are provided in the main text. □

References

An, Yu, "Competing with Inventory in Dealership Markets," *Working Paper*, 2019, *Johns Hopkins Carey Business School.*

———, **Wei Li, and Zhaogang Song**, "TBA Trading and Security Issuance in the Agency MBS Market," *working paper*, 2022.

Asquith, Paul, Thomas R. Covert, and Parag A Pathak, "The Effects of Mandatory Transparency in Financial Market Design: Evidence from the Corporate Bond Market," *Working Paper*, 2019.

Bank for International Settlements, "Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools," *Basel Committee on Banking Supervision*, 2013.

Bessembinder, Hendrik, Chester Spatt, and Kumar Venkataraman, "A Survey of the Microstructure of Fixed-Income Markets," *Journal of Financial and Quantitative Analysis*, 1 2019.

———, **William Maxwell, and Kumar Venkataraman**, "Trading activity and transaction costs in structured credit products," *Financial Analysts Journal*, 2013, *69(6)*, 55–68.

Chen, Jiakai, Haoyang Liu, Asani Sarkar, and Zhaogang Song, "Dealers and the Dealer of Last Resort: Evidence from the Agenyc MBS Market in the COVID-19 Crisis," *Federal Reserve Bank of New York Staff Reports*, 2023, (933).

DeMarzo, Peter M., "The Pooling and Tranching of Securities: A Model of Informed Intermediation," *Review of Financial Studies*, 2005, *18* (1), 1–35.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, "The value of trading relations in turbulent times," *Journal of Financial Economics*, 2017, *124(2)*, 266–284.

Downing, Christopher, Dwight Jaffee, and Nancy Wallace, "Is the Market for Mortgage Backed Securities a Market for Lemons?," *Review of Financial Studies*, 2009, *22–7*, 2457–2494.

Duffie, Darrell, *Dark Markets: Asset Pricing and Information Transmission in Over-the-Counter Markets*, New Jersey: Princeton University Press, 2012.

Eisfeldt, Andrea, Hanno Lustig, and Lei Zhang, "Complex Asset Markets," *Working Paper*, 2019.

Fleming, Michael J., Bruce Mizrach, and Giang Nguyen, "The microstructure of a U.S. Treasury ECN: The BrokerTec platform," *Journal of Financial Markets*, 2018, *40*, 2–22.

French, Kenneth R. and Robert E. McCormick, "Sealed Bids, Sunk Costs, and the Process of Competition," *The Journal of Business*, 1984, *57* (4), 417–441.

Fusari, Nicola, Wei Li, Haoyang Liu, and Zhaogang Song, "Asset Pricing with Cohort-Based Trading in MBS Markets," *Journal of Finance (Forthcoming)*, 2022.

Gao, Pengjie, Paul Schultz, and Zhaogang Song, "Liquidity in a Market for Unique Assets: Specified Pool and TBA Trading in the Mortgage Backed Securities Market," *Journal of Finance*, 2017, *72-3*, 1119–1170.

______, ______, and ______, "Trading Methods and Trading Costs for Agency Mortgage Backed Securities," *Journal of Investment Management*, 2018, *forthcoming.*

Glaeser, Edward and Hedi Kallal, "Thin Markets, Asymmetric Information, and Mortgage-Backed Securities," *Journal of Financial Intermediation*, 1997, *6*, 64–86.

Gorton, Gary and George Pennacchi, "Financial Intermediaries and Liquidity Creation," *Journal of Finance*, 1990, *45* (1), 49–71.

Hirshleifer, Jack, "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 1971, *61* (4), 561–574.

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

Li, Jessica, Maureen O'Hara, Andreas Rapp, and Xing (Alex) Zhou, "Bond Market Illiquidity: Is Portfolio Trading the Solution?," *working paper*, 2023.

Liu, Haoyang, Zhaogang Song, and James Vickery, "Fragmentation, Liquidity, and Pricing in MBS Markets," *working paper*, 2020.

Milbradt, Konstantin, "Asset heterogeneity in Over-The-Counter markets," *Working Paper*, 2018.

Schultz, Paul and Zhaogang Song, "Transparency and dealer networks: Evidence from the initiation of post-trade reporting in the mortgage backed security market," *Journal of Financial Economics*, 2019, *133* (1), 113–133.

Song, Zhaogang and Haoxiang Zhu, "Quantitative easing auctions of Treasury bonds," *Journal of Financial Economics*, 2018, *128* (1), 103 – 124.

Spatt, Chester, "Frictions in the Bond Market," *Keynote Speech: Second MTS Conference on Financial Markets*, 2004.

Subrahmanyam, Avanidhar, "A Theory of Trading in Stock Index Futures," *Review of Financial Studies*, 1991, *4* (1), 17–51.

Üslü, Semih and Güner Velioglu, "Liquidity in the Cross Section of OTC Assets," *Working Paper*, 2019.

Vayanos, Dimitri and Jiang Wang, "Chapter 19 - Market Liquidity—Theory and Empirical Evidence," *Handbook of the Economics of Finance*, 2013, *2*, 1289 – 1361.

——— **and Pierre-Olivier Weill**, "A search-based theory of the on-the-run phenomenon," *Journal of Finance*, 2008, *63*, 1361–1398.

——— **and Tan Wang**, "Search and endogenous concentration of liquidity in asset markets," *Journal of Economic Theory*, 2007, *136* (1), 66 – 104.

Vickery, James and Joshua Wright, "TBA Trading and Liquidity in the Agency MBS Market," *Federal Reserve Bank of New York Economic Policy Review*, 2011, *19*.

Weill, Pierre-Olivier, "Liquidity premia in dynamic bargaining markets," *Journal of Economic Theory*, 2008, *140* (1), 66–96.

———, "The Search Theory of Over-the-Counter Markets," *Annual Review of Economics*, 2020, *12* (1), 747–773.

Internet Appendix

In this Internet Appendix, we briefly discuss a few extensions to the main model.

A.1 Design of QC Contracts

In this section, we study the design of QC contracts based on our theoretical framework. The optimal design of QC contracts in general has no closed-form solutions because of the complications resulting from asset heterogeneity. Traders' venue choices, for example, depend on the distribution of asset values (as Eq. (42) shows). Hence, we focus on a specific issue of market design motivated directly by market reforms in practice—setting the eligibility criterion of existing QC contracts.

In particular, to support the mortgage market, the Economics Stimulus Act enacted in February 2008 allowed Fannie Mae and Freddie Mac to purchase and securitize high-balance loans (known as super-conforming loans). However, the SIFMA announced to prohibit MBS backed by these loans from being delivered into TBA contracts immediately. Later, in August 2008, SIFMA announced that MBS of super-conforming loans would be TBA-eligible but imposed that these loans could represent at most 10 percent of a TBA pool (Vickery and Wright, 2011). Because high-balance mortgages tend to have high prepayment risks and low values (Fusari et al., 2022), restricting their TBA-eligibility effectively increased the TBA-eligibility threshold, as represented by $\underline{v}$ in our model. Below, we study the effects of varying $\underline{v}$, which changes the set of deliverable assets.

Based on the parallel-trading market equilibrium in Section 4, varying $\underline{v}$ changes the proportion of assets traded on the QC market q (defined in Eq. (43)) both by directly determining the set of assets that are excluded from QC trading $[v_{\min}, \underline{v})$ and, as Eq. (42) shows, by indirectly affecting the fraction of QC-eligible assets that would be traded via QC contracts as captured by $[\underline{v}, \bar{v}]$. The direct effect of raising the QC-eligible threshold $\underline{v}$ is unambiguous: it shrinks the set of QC-eligible assets. The indirect effect of raising $\underline{v}$, however, is ambiguous. On the one hand, raising $\underline{v}$ could raise $\bar{v}$ because doing so excludes low-value assets from QC trading and mitigates the CTD discount of QC pricing, which could attract sellers of high-value assets to the QC market. On the other hand, raising $\underline{v}$ could lower $\bar{v}$ because it reduces the maximum liquidity externality benefits—fewer assets could be traded together on the QC market—and could drive sellers of high-value assets to AS trading.

Because raising $\underline{v}$ may increase or decrease $\bar{v}$, its effects on the proportion of assets traded on the QC market q and on market liquidity are ambiguous.[22] Nonetheless, SIFMA's restrictions on TBA-eligibility of high-balance loans mentioned above imply that regulators believed that lowering in $\underline{v}$ would likely shrink the proportion of QC-traded assets and hurt liquidity. In fact, SIFMA Vice Chairman Thomas Hamilton explained, in his testimony to the House Committee on Financial Services in May 2008, that making high-balance loans TBA-eligible "would ... drive trading into the specified pool market" and "negatively impact the liquidity of the product (TBA)."[23] Accordingly, we focus on the situation in which raising $\underline{v}$ increases q.

Fig. IA.4 illustrates this situation. In particular, with the QC-eligible threshold (exogenously) increased from $\underline{v}_1$ to $\underline{v}_2$ and the upper bound of QC asset values (endogenously) increase from $\bar{v}_1$ to $\bar{v}_2$, assets can be partitioned into four subsets: (1) both the least valuable assets in group ① and the most valuable assets in group ⑤ remain on AS submarkets, (2) assets in group ② switch from the QC market to AS submarkets, (3) assets in group ③ remain on the QC market, and (4) assets in group ④ switch from AS submarkets to the QC market.[24] The following result presents the effects on liquidity of these four subsets of assets.

Corollary IA.9. *Consider the situation in which an increase of $\underline{v}$ leads to an increase in q. In this situation, liquidity (1) weakly declines for assets that remain on AS submarkets, (2) weakly declines for assets that switch from the QC market to AS submarkets, (3) weakly improves for assets that switch from AS submarkets to the QC market, and (4) may improve or decline for assets that remain on the QC market. The average liquidity across assets and the total trading volume both increase.*

Proof of Corollary IA.9. When q increases, $\mu(Sq)$ increases and, according to Eq. (44), $\widetilde{b}^*$ decreases. It follows that $\widetilde{\pi}^b$ increases and $\widetilde{\pi}^s_{\text{as}}$ decreases. In addition, Eq. (49) implies

[22]Through numerical analysis, we have verified that both scenarios may occur.

[23]Available at https://www.sifma.org/resources/submissions/testimony-on-behalf-of-sifma-at-hcfs-hearing-on-conforming-loan-limit-increase-and-impact-on-homebuyers-and-housing-market/.

[24]By contrast, as we show in Section 4, only two subsets of assets exist when QC trading is introduced to a market with only AS trading: assets that remain on AS submarkets and assets that switch from AS submarkets to the QC market.

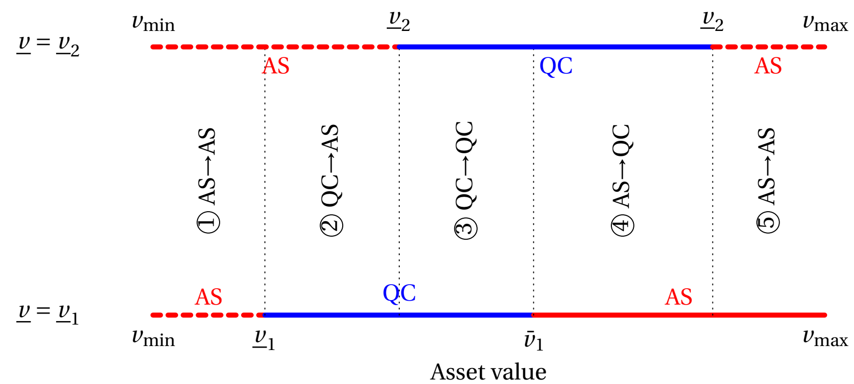


Figure IA.4. Illustration of the situation in which increasing $\underline{v}$ leads to increasing q.

that $\widetilde{b}_0$ decreases, so the mass of participating buyers $B-\widetilde{b}_0$ increases. Thus, total trading volume $\widetilde{V} = \widetilde{\pi}^b(B-\widetilde{b}_0)$ increases and social welfare $\widetilde{\Omega} = (B-\widetilde{b}_0)(\widetilde{\pi}^b\delta - c)$ increases. □

That is, the effect on liquidity varies across assets. First, liquidity deteriorates for assets that remain on AS submarkets because, as Eq. (44) shows, the mass of buyers on each AS submarket $\widetilde{b}^*$ weakly decreases with q. Second, liquidity deteriorates for assets that switch from QC trading to AS trading and improves for assets that switch from AS trading to QC trading. Consider an asset that switches from QC trading to AS trading when $\underline{v}$ increases. Based on Eq. (58), this asset's liquidity level before the switch is above π^s (based on the second inequality) and its liquidity level after the switch is below π^s (based on the first inequality), where π^s represents the liquidity level in the AS-only equilibrium. Thus, this asset becomes less liquid. Similarly, assets that switch from AS trading to QC trading when $\underline{v}$ increases must become more liquid. Third, liquidity may improve or deteriorate for assets that remain on the QC market. In particular, as q increases, the buyer-to-seller ratio on the QC market $B_{\text{qc}}/S_{\text{qc}}$ does not monotonically increase because the ratio equals B/S when the QC market consists of one type of assets ($q = 1/S$) or all types of assets ($q = 1$). Thus, although $B_{\text{qc}}/S_{\text{qc}}$ increases in certain ranges of q, it decreases in other ranges, which may result in worse liquidity on the QC market.

Although its liquidity effects vary across assets, we show that increasing $\underline{v}$ to raise q always increase the overall liquidity (measured by total trading volume $\widetilde{V}$) because, intuitively, it reduces the proportion of assets traded on the more frictional AS submarkets. Thus, if increasing the eligibility threshold $\underline{v}$ indeed increases q, then SIFMA's decision is justified: limiting super-conforming loans in TBA pools improves liquidity for TBA contracts and such improvement dominates negative liquidity effects on SP contracts.[25]

A.2 Participation Cost

Next, we present the effects of lowering the participation cost c in the parallel-trading market as follows.

Corollary IA.10. *Suppose that Assumption 1 holds. After c declines, sellers' venue choice ($\bar{v}$) does not change. Moreover,*

- *if all buyers participate in trading ($\widetilde{b}_0 = 0$) before c declines, then after c declines, buyer mass and asset liquidity on every venue remain unchanged;*
- *if not all buyers participate in trading ($\widetilde{b}_0 > 0$) before c declines, then after c declines, buyer mass and asset liquidity on every venue strictly increase.*

Proof of Corollary IA.10. Lowering c increases $\bar{b} = (\lambda\rho\delta/c)^{\frac{2}{1-\theta}}$ but as Eq. (42) implies, it does not affect $\bar{v}$, so q, S_{as}, and S_{qc} stay the same.

If $\widetilde{b}_0 = 0$, then $B \le \bar{b}(S + \mu(Sq))$. Thus, $\widetilde{b}^* = B/(S + \mu(Sq))$ is unaffected by c, which impiles that $\widetilde{\pi}^s_{\text{as}}$ and $\widetilde{\pi}^s_{\text{qc}}$ stay the same.

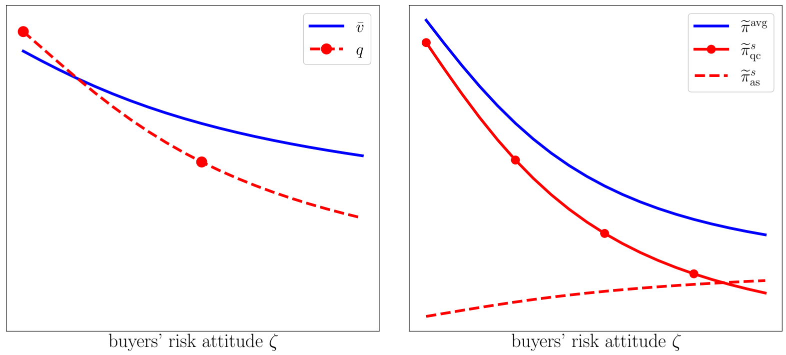
If $\widetilde{b}_0 > 0$, then $B > \bar{b}(S + \mu(Sq))$. Hence, lowering c increases $\widetilde{b}^* = \min\left\{\frac{B}{S+\mu(Sq)}, \bar{b}\right\}$ and reduces $\widetilde{b}_0 = \max\{B - (S + \mu(Sq))\bar{b}, 0\}$. Consequently, $\widetilde{\pi}^s_{\text{qc}}$ and $\widetilde{\pi}^s_{\text{as}}$ both increase. □

First, reducing buyer participation cost c does not affect sellers' venue choices. Because all venues share the same c, if additional buyers participate after c declines, they would join all venues proportionally so that they are indifferent between venues. Thus, the ratios of asset liquidity across venues remain unchanged, so no sellers switch venues.

Second, if all buyers already participate ($\widetilde{b}_0 = 0$), reducing c cannot attract additional buyers to participate. In this situation, the masses of buyers and sellers on every venue remains unchanged, so asset liquidity is unaffected.

Third, if, before c declines, some buyers were not participating ($\widetilde{b}_0 > 0$) and earn zero profit, then all buyers earn zero profits. After c declines, buyers would earn more

[25] Note that in theory raising $\underline{v}$ could lower q; in this situation, the overall liquidity would decline. Thus, it is necessary to empirically verify the premise that limiting super-conforming loans raises q.



(a) QC market properties   **(b)** Trading volume

Figure IA.5. Impact of buyers' risk attitude ζ.

expected profits by participating in every venue; thus, on every venue, additional buyers join, raising the buyer-to-seller ratio and improving asset liquidity. As a result, every seller earns more expected profit because she is more likely to sell her asset. Intuitively, reducing buyer participation cost c benefit sellers if it increases the level of competition between participating buyers. In particular, if some buyers still do not participate after c declines, then all buyers still earn zero profits and all gains resulting from the reduction in c would be extracted by sellers.

A.3 Buyers' Risk Attitude

In this section, we examine effects of allowing buyers to be risk-averse. In particular, we assume that buyers in the QC market bid $\hat{P}_{\text{qc}}$ and accept the ask price $\hat{P}_{\text{qc}} + \delta$, where

$$\hat{P}_{\text{qc}} = \zeta \cdot \underline{v} + (1-\zeta)\,\mathbf{E}[v_k | k \in \widetilde{\mathscr{M}}_{\text{qc}}]. \tag{IA.14}$$

When buyers' risk attitude $\zeta = 0$, buyers are risk-neutral; when $\zeta = 1$, buyers are ambiguity-averse and bid the value of the cheapest QC asset $\underline{v}$. Fig. IA.5 shows our numerical results. As $\zeta \uparrow$, buyers bid lower prices, driving some sellers off the QC market ($\bar{v} \downarrow$). As a result, buyers are less likely to trade on the QC market, so some of them switch to AS submarkets, thereby improving AS market asset liquidity $\widetilde{\pi}^{s}_{\text{as}}$. Because fewer traders use QC trading, the total trading volume $\widetilde{V}$ declines.