Asset Heterogeneity, Market Fragmentation, and Quasi-Consolidated Trading^{*}

Wei Li[†] Zhaogang Song[‡]

December 18, 2024

Abstract

Significant asset heterogeneity characterizes many over-the-counter (OTC) markets. We first introduce a benchmark model that includes only standard asset-specific (AS) trading and demonstrate that asset heterogeneity fragments the market, thereby diminishing the positive network effects on liquidity. Next, we model a parallel-trading market that also allows quasiconsolidated (QC) trading—where assets with varying values trade at a uniform price—and demonstrate that this parallel-trading market results in higher total trading volume and greater social welfare. However, QC trading's uniform pricing creates a "cheapest-to-deliver" effect, hurting liquidity for sellers who do not participate in QC trading. Additionally, these sellers, along with some sellers who participate in QC trading, earn lower profits than they would in the AS-only market. Our model provides a foundation for analyzing liquidity and market design in OTC markets with 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

[†]Zicklin School of Business, Baruch College. E-mail:wei.li@baruch.cuny.edu.

^{*}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, Jiakai Chen, 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, the 2023 EFA meeting, and the 2024 AREUEA-ASSA meeting.

[‡]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 overthe-counter (OTC) markets, has advanced rapidly, particularly following the 2008 financial crisis. Theoretical studies have examined key features of OTC markets, such as search frictions, bargaining, dealer intermediation, network structures, and market opacity (see Duffie (2012) and Weill (2020) for surveys). In this paper, we focus on another salient aspect of OTC markets: asset heterogeneity.

Substantial asset heterogeneity characterizes many key OTC markets, including U.S. fixed-income markets. For instance, as of 2022, the corporate bond market featured over 100,000 bonds issued by approximately 5,600 firms. Similarly, by the end of 2017, the municipal bond market comprised more than 1.5 million bonds issued by around 50,000 entities (Bessembinder, Spatt and Venkataraman, 2019). In the agency mortgage-backed securities (MBS) market, over 800,000 securities backed by mortgages with diverse borrower, loan, and lender characteristics were outstanding as of 2019.

This high degree of asset heterogeneity is widely regarded as a factor impairing OTC market liquidity. For example, Bessembinder et al. (2019) argue 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." Supporting this view, the average daily trading volume in 2019 was about \$34 billion for corporate bonds and \$12 billion for municipal bonds, yielding annual turnover rates of 1.1-1.3. In contrast, the U.S. Treasury market, characterized by relatively low heterogeneity (with around 400 securities outstanding), had a daily trading volume of approximately \$590 billion and an annual turnover of 12.8.

The agency MBS market, however, stands out as an intriguing exception. Despite its pronounced asset heterogeneity, it is considered "one of the most liquid fixed-income markets in the world, with trading volumes typically in the trillions of dollars per year."¹ Its primary trading mechanism, the to-be-announced (TBA) contract, allows a cohort of

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

heterogeneous MBS to trade at a *uniform* price. The TBA market achieves an average daily volume of approximately \$230 billion and an annual turnover of 10.7, comparable to the U.S. Treasury market (Vickery and Wright, 2011; Gao, Schultz and Song, 2017). While one might attribute this exceptional liquidity to the government backing of agency MBS, this explanation proves insufficient. Specifically, trading agency MBS through standard individual-security-based contracts, known as specified pool (SP) contracts—similar to those used for corporate and municipal bonds—is much less active, with an annual turnover of just 0.8, comparable to that of corporate and municipal bonds. This suggests that TBA trading, rather than government guarantees, is the primary driver of the exceptional liquidity in agency MBS.²

How does asset heterogeneity impair OTC market liquidity? How does TBA trading enhance liquidity in the presence of asset heterogeneity? Do these effects vary across assets and investors? Addressing these questions is crucial not only for understanding agency MBS markets—which finance the majority of residential mortgages³—but also for improving the design of other OTC markets. Indeed, it has been conjectured that introducing a TBA-like trading mechanism could enhance the liquidity of corporate and municipal bonds.⁴ In this paper, we build a theoretical framework to address these questions, laying a foundation for analyzing liquidity and the design of OTC markets involving heterogeneous assets.

In the first part of our analysis, we develop a benchmark model that illustrates how asset heterogeneity causes market fragmentation and impairs OTC market liquidity by limiting the positive network externalities on liquidity.

Specifically, in the benchmark model, assets with heterogeneous values are traded through standard contracts that fully specify the assets; we refer to these as asset-specific (AS) contracts. In this AS-only market, asset heterogeneity causes fragmentation be-

²Liquidity patterns are similar when measured using trading costs; see Section 2 for further details.

³As of Q3 2020, 63% of the \$11.5 trillion outstanding residential mortgage debt was securitized into agency MBS, per Urban Institute reports.

⁴ For instance, Spatt (2004) notes, "The analogy to the mortgage markets is instructive. Trading instruments based upon their main characteristics ... may be helpful and narrow the spreads." Similarly, 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."

cause (1) a seller can sell only her own asset, and (2) while buyers face no restrictions on which assets they can purchase,⁵ each buyer incurs a cost for participating in the trading of any given asset.⁶ For simplicity, we focus on the case where participation costs are so high that each buyer can trade at most one asset. In equilibrium, the AS-only market fragments into multiple segmented AS submarkets, each with a distinct subset of traders using an AS contract for a specific asset.

The trading follows a standard static search-and-matching process. Specifically, on each trading venue (e.g., an AS submarket as defined above), buyers and sellers are randomly matched, and then a trade occurs between every matched buyer-seller pair. Accordingly, we define the *liquidity of an asset* as the probability of successfully selling it. We assume that the matching function exhibits increasing returns to scale, meaning that pooling more traders in the same trading venue enhances liquidity. This reflects the positive network externalities on market liquidity commonly observed in OTC markets.(Vayanos and Weill, 2008; Weill, 2020). We show that asset heterogeneity leads to market fragmentation, reducing the number of traders at each venue, which weakens network externalities and impairs liquidity.

In the second part of our analysis, we extend the benchmark setup to include a TBAlike contract alongside the standard SP contract. Specifically, the TBA-like contract allows sellers to deliver any asset from a predefined cohort at a uniform price. In contrast to AS contracts, which set asset-specific prices, a TBA-like contract establishes a single price for all deliverable assets within a cohort of heterogeneous assets. We refer to this TBA-like contract as a quasi-consolidated (QC) contract, as buyers pay the same price but may receive different assets. By comparison, security baskets and exchange-traded funds are *fully* consolidated, as they combine multiple assets into a single traded security.

⁵In practice, many investors do not favor specific assets within a class. For example, under Basel III's liquidity coverage ratio (LCR) requirements, all MBS guaranteed by Fannie Mae and Freddie Mac are treated as level 2A assets and receive the same haircut in computing the amount of "high-quality liquid assets" (Bank for International Settlements, 2013). Similarly, as noted by Spatt (2004), "typically, buyers are not focused upon particular securities but instead are interested in purchasing a security with certain characteristics."

⁶This participation cost, which is common in the literature (Vayanos and Wang, 2013), arises naturally from asset heterogeneity. For example, when assets differ in value, buyers must collect data, forecast cash flows, and run models to analyze and value an asset before bidding.

When both AS trading and QC trading are available, a seller owning a QC-eligible asset (i.e., an asset deliverable under the QC contract) can choose to sell it through either the AS contract specific to that asset or the QC contract. Similarly, a buyer can opt for the QC contract or any AS contract. As in the benchmark setup, we assume that buyers incur a cost to participate in the trading of each contract (either the QC contract or an AS contract) and still consider the high-participation-cost scenario, where each buyer trades through at most one contract. In equilibrium, the market fragments into one QC market and multiple AS submarkets, each comprising a distinct subset of traders using a specific contract.⁷

In the parallel-trading equilibrium, a "cheapest-to-deliver" (CTD) practice emerges in QC trading: since all eligible assets are priced uniformly, low-value assets are more likely to be delivered under the QC contract. This CTD effect discourages sellers from choosing QC trading if it is strictly less liquid than AS trading. Consequently, in equilibrium, QC trading is weakly more liquid than AS trading. Sellers participating in QC trading benefit from better liquidity but face price discounts if their assets are more valuable than the uniform QC price. Conversely, sellers opting for AS trading experience worse liquidity but receive asset-specific prices. As a result, in equilibrium, sellers of QC-eligible assets choose AS trading only if their asset values exceed an endogenous threshold.

We then contrast the parallel-trading market equilibrium with the AS-only market equilibrium to assess the impact of QC trading on market liquidity, traders' profits, and social welfare.

In terms of market liquidity, we show that, compared to the AS-only equilibrium, liquidity improves for assets in the QC market but declines for those in AS submarkets under the parallel-trading equilibrium. Specifically, liquidity improves in the QC market for two reasons. First, for any given buyer-to-seller ratio, QC trading enhances liquidity through network externalities generated by concentrating more traders in a single venue. Second, QC trading disproportionately attracts more buyers than sellers, increasing the buyer-to-seller ratio in the QC market and further boosting the likelihood of successful asset sales. This disproportionate adoption arises from differences in partic-

⁷Degenerate equilibria, where only AS or QC trading occurs, exist and are straightforward to analyze, but do not align with practices observed in markets like the agency MBS market.

ipation costs: while buyers face uniform costs across venues, QC trading is more costly for sellers of high-value assets due to CTD pricing. Conversely, liquidity declines in AS submarkets for a similar reason: QC trading disproportionately attracts buyers, lowering buyer-to-seller ratios in AS submarkets and reducing liquidity for assets traded through AS contracts.

Although QC trading leads to varying liquidity effects across assets, it enhances *overall* market liquidity, measured by the average trading probability of all assets.⁸ Intuitively, by enabling multiple assets to be traded together in the parallel-trading equilibrium, QC trading partially "defragments" the benchmark AS-only market. This defragmentation, combined with the network externalities captured by the increasingreturns-to-scale matching function, reduces overall trading frictions and improves aggregate market liquidity.

We then compare traders' profits across the AS-only and parallel-trading equilibria. First, we show that, compared to the AS-only equilibrium, buyers are more likely to trade successfully and, as a result, earn higher expected profits in the parallel-trading equilibrium. Specifically, in the parallel-trading equilibrium, buyers participating in QC trading benefit from returns to scale in liquidity, while those opting for AS trading benefit from lower buyer-to-seller ratios relative to the AS-only equilibrium. Furthermore, these higher buyer profits lead to greater buyer participation in the parallel-trading equilibrium compared to the AS-only equilibrium. However, sellers using AS trading in the parallel-trading equilibrium earn lower expected profits because they receive the same prices but face a reduced likelihood of trading. Finally, compared to the AS-only equilibrium, sellers choosing QC trading in the parallel-trading equilibrium also earn lower expected profits if their assets are more valuable than an endogenous break-even level. Although these sellers opt for QC trading, they are worse off because QC trading reduces liquidity in AS submarkets, diminishing the value of their outside options to sell there.

Additionally, compared to the AS-only equilibrium, although not all traders earn higher profits in the parallel-trading equilibrium, overall social welfare—defined as the total expected profits of all traders—improves. Specifically, the total profit of all buyers increases because more buyers participate and each participating buyer earns higher

⁸In our model, this probability is proportional to total trading volume and turnover, as the set of assets is fixed.

expected profits. The total profit of all sellers also rises because total trading volume increases, while the average profit per trade for sellers remains unchanged.

Finally, we extend the model to demonstrate that our key results remain valid when we relax simplifying assumptions, such as homogeneous participation costs and uniform trading gains across assets. Additionally, we analyze the effects of varying QCeligibility criteria, inspired by market practices like excluding MBS with specific features (e.g., high loan balances or loan-to-value ratios) from TBA delivery.

Most existing theoretical studies in the OTC market literature model the trading of a single asset, whereas asset heterogeneity is a key feature of our model. Among the few studies that consider markets with multiple assets, Vayanos and Wang (2007) and Vayanos and Weill (2008) also feature endogenous market fragmentation, similar to our approach.⁹ These studies, however, focus on explaining why assets with the *same* value and can trade at different prices. In contrast, we model assets with *heterogeneous* values and focus on analyzing the effects of a trading mechanism that can pool the trading of heterogeneous assets, which could potentially be applied in various OTC markets. To the best of our knowledge, our model is the first to examine how QC trading affects the liquidity of OTC markets for assets with heterogeneous values.¹⁰

2 Background and Motivation

In this section, we provide a brief overview of the institutional background of OTC markets for U.S. fixed-income assets, emphasizing the prevalence of asset heterogeneity and the potential role of QC trading in enhancing market liquidity. We then contrast the key economic mechanism in our model with existing (mostly informal) explanations for the liquidity of QC trading.

⁹Other studies include Weill (2008), Milbradt (2018), An (2019) and Üslü and Velioglu (2019).

¹⁰Our paper also contributes to the largely empirical literature on the agency MBS market structure and liquidity, including works by Bessembinder, Maxwell and Venkataraman (2013), Downing, Jaffee and Wallace (2009), Gao, Schultz and Song (2018), and Schultz and Song (2019), among others.

2.1 Institutional Background

U.S. fixed-income markets are a major source of financing for the economy. As of 2019, according to SIFMA, the outstanding balances for Treasury securities, agency MBS, corporate bonds, and municipal bonds were approximately \$16.7 trillion, \$7.7 trillion, \$9.6 trillion, and \$3.9 trillion, respectively (see the first column of Table 1).¹¹ The trading of these securities occurs primarily in opaque, decentralized, and dealer-intermediated OTC markets, rather than on centralized exchanges.¹²

A salient feature of these fixed-income markets is significant asset heterogeneity. For instance, as of June 2022, there were 105,132 corporate bonds outstanding, issued by 5,607 firms (see the second column of Table 1). Additionally, according to Bessembinder et al. (2019), there were over 1.5 million municipal bonds issued by approximately 50,000 issuers as of December 2017. This considerable asset heterogeneity is widely believed to reduce market liquidity, as discussed in the Introduction. Indeed, the average daily trading volumes for corporate and municipal bonds are about \$34 billion and \$12 billion, respectively, yielding annual turnovers of 1.3 and 1.1, respectively. In contrast, the U.S. Treasury market, characterized by very low heterogeneity with only 404 securities (as of 2019), has an average daily trading volume of approximately \$594 billion, resulting in an annual turnover of about 12.8. Trading costs follow a similar pattern: they are around 80-100 basis points (bps) for corporate and municipal bonds (Di Maggio, Kermani and Song, 2017; Asquith, Covert and Pathak, 2019; Li and Schürhoff, 2019), but only a few 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, its trading remains very active and liquid. Specifically, while over 800,000 agency MBS are outstanding (as of 2019), the average daily trading volume is approximately \$246 billion, resulting in an annual turnover of about 11.5. This turnover is comparable to that of Treasury securities (see the fourth row of Table 1).

¹¹Other important but smaller fixed-income markets include non-agency MBS (\$1.4 trillion), federal agency securities (\$1.8 trillion), and asset-backed securities (\$1.8 trillion).

¹²A small fraction of fixed-income trading occurs on centralized limit-order books. For instance, in the U.S. Treasury market, the inter-dealer segment of on-the-run securities trades through a centralized limit-order book operated by BrokerTec (Fleming, Mizrach and Nguyen, 2018). Additionally, about half of the inter-dealer trades of agency MBS are executed on a centralized limit-order book run by TradeWeb (Schultz and Song, 2019).

Markets	Outstanding	Number of	Trading	
	(\$ tn)	securities	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

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

This table provides aggregate summaries of U.S. fixed-income markets, including municipal bonds, corporate bonds, Treasury securities, and agency MBS. The first column reports the dollar outstanding amounts as of 2019 based on the SIFMA reports. The second column details the number of securities outstanding: for municipal bonds (as of 2017), data is taken from Bessembinder et al. (2019); for corporate bonds (as of 2022), from the Mergent Fixed Income Securities Database; for Treasury securities (as of 2019), from 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, from eMBS data. The third column reports the average daily trading volume in 2019 along with the related annual turnover (calculated as total trading volume in 2019 divided by outstanding amount as of 2019), based on the SIFMA reports. The fourth column provides 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.

The trading of agency MBS occurs primarily through TBA forward contracts. In particular, a TBA contract does not specify the specific MBS to be delivered at settlement; instead, it specifies characteristics such as the issuer (e.g., Fannie Mae), the type of MBS (e.g., 30-year fixed-rate), and the security coupon rate (e.g., 4%) (Gao et al., 2017).¹³ TBA trading has been shown to significantly enhance the trading activity and liquidity of agency MBS (Gao et al., 2017; Bessembinder et al., 2019). In fact, the average daily TBA trading volume is \$229 billion, accounting for approximately 93% of the total daily trading volume of all agency MBS, resulting in an annual turnover of 10.7. Additionally,

¹³The specific MBS to be delivered under a TBA contract is identified only two days before the settlement date. Details of 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.

the TBA trading cost is about 2-5 basis points (bps). In contrast, the SP trading of agency MBS—similar to the standard individual-security trading of corporate and municipal bonds—is much less liquid. Its average daily trading volume is an order of magnitude lower than TBA trading, with trading costs reaching up to 60 bps (Gao et al., 2017), comparable to those of corporate and municipal bonds. This stark contrast suggests that it is primarily the TBA trading mechanism, rather than the government backing of agency MBS, that contributes to the remarkable liquidity of the agency MBS market.

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

2.2 Discussion of the Economic Mechanism

The institutional background and anecdotal evidence outlined above suggest that asset heterogeneity hinders liquidity in OTC markets, whereas the TBA-like cohort-based trading mechanism mitigates frictions caused by asset heterogeneity and enhances market liquidity. Before presenting our formal model of OTC market liquidity under asset heterogeneity, it is helpful to discuss and clarify several informal arguments on how TBA trading improves liquidity. As noted in the Introduction, no theoretical models or formal analyses have previously addressed this topic.

First, motivated by the feature of the TBA contract in which the specific MBS to be delivered is not disclosed to buyers at the time of trade, several studies suggest that TBA trading enhances liquidity by limiting information disclosure and mitigating adverse selection.¹⁴ In this context, Vickery and Wright (2011) argue that TBA trading is analogous to the De Beers diamond auction, stating that "the limits on information disclosure in-

¹⁴The notion that limiting information disclosure can reduce adverse selection and improve liquidity is reminiscent of the "Hirshleifer effect" (Hirshleifer, 1971).

herent 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." However, as French and McCormick (1984) point 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, where sellers are known to engage in the "cheapest-to-deliver" practice; for formal evidence, see Fusari, Li, Liu and Song (2022).

Relatedly, Glaeser and Kallal (1997) develop a formal model showing that the liquidity of agency MBS can be improved when initial issuers limit the disclosure of MBS characteristics, thereby preventing dealers from analyzing these characteristics and gaining a substantial information advantage over investors. However, their model focuses on the liquidity of *all* agency MBS trading and does not explain why TBA trading is more liquid than SP trading. In their informal argument based on the model, they suggest 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—specifically, which MBS will be delivered by issuers—will eventually be revealed to dealers after delivery. As a result, TBA contracts *cannot* reduce the information advantage of dealers over investors.¹⁵

Second, in addition to using TBA contracts to trade MBS, as we model, investors can also use TBA contracts for hedging because they are forward contracts. For instance, mortgage lenders often use TBA contracts to hedge interest rate risks (Vickery and Wright, 2011), and dealers use them to hedge their MBS inventory (Chen, Liu, Sarkar and Song, 2023). These hedging activities can further increase trading volumes and enhance the liquidity of TBA contracts. Nonetheless, while trades driven purely by hedging

¹⁵Glaeser and Kallal (1997) also propose another view, suggesting that only "the 'worst' possible mortgage bundle" is traded in the TBA market, which they assume has a known value to all traders and is thus perfectly liquid. This view, however, is inconsistent with market practice, where "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). Indeed, 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.

are typically settled by offsetting trades, a significant proportion of TBA contracts are settled by physical delivery of MBS. This suggests that investors use TBA contracts not only for hedging but also as an important mechanism to buy and sell MBS. Our study focuses on this crucial trading function of TBA contracts.

Third, we take asset heterogeneity as given and focus on analyzing the trading mechanism. While security issuers can 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—our study differs by analyzing how a specific trading mechanism can enhance liquidity without altering the securities being issued.¹⁶

3 Asset Heterogeneity and Market Fragmentation

In this section, we develop a benchmark model featuring only standard asset-specific (AS) trading and demonstrate how asset heterogeneity leads to market fragmentation, thereby restricting market liquidity.

3.1 Setup

Assets are traded bilaterally between a continuum of risk-neutral buyers and sellers.¹⁷ We assume there are B buyers and S assets, denoted by

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

For simplicity, we normalize the mass of each asset to 1, so the total mass of assets equals *S*. Each seller owns one share of a specific asset, meaning the mass of sellers who own each asset also equals 1 and the total mass of sellers equals *S*. A buyer may choose any

¹⁶A related trading mechanism, called portfolio trading, has gained popularity in recent years (Li, O'Hara, Rapp and Zhou, 2023). In portfolio trading, an investor submits a *single order* consisting of a basket of bonds to multiple dealers and executes the order with the dealer who can fill the entire order. An important distinction is that a seller needs to have *all* the bonds in the basket to use portfolio trading, whereas a seller with *any* eligible MBS can engage in TBA trading.

¹⁷In Internet Appendix IA.2, we model risk-averse buyers.

asset but can purchase up to one share of this asset. Since buyers may opt out of trading, the mass of buyers *participating* in the market may be less than *B*.

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

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

Thus, a trade between a seller and a buyer generates a trading gain of $\delta > 0$.¹⁸

Trades occur through standard AS contracts, where each contract allows the delivery of one asset. For clarity, we define traders as being on the same trading venue (an AS submarket) if they trade using the same contract. Two types of trading frictions arise, both linked to asset heterogeneity:

The first type of friction, consistent with many OTC models, is search friction. For simplicity, we use a static search-and-matching framework to model this friction. Buyers and sellers on any trading venue are randomly chosen and matched before trading can occur; unmatched agents cannot trade. Formally, if sellers with mass m_s and buyers with mass m_b participate in a trading venue, the expected number 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 the exogenous parameter λ measures matching efficiency, and θ captures the liquidity benefit from pooling multiple traders on one venue.

Particularly, we measure the liquidity level of a trading venue by the probability that a seller on this venue is matched with a buyer. Based on Eq. (3), this probability equals

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

Thus, the liquidity level of a trading venue increases with both the buyer-to-seller ratio m_b/m_s and the mass of buyers m_b on this venue. Similarly, based on Eq. (3), the

¹⁸In Appendix B.2, we demonstrate that our main results on asset liquidity remain valid when trading gains vary across assets.

probability that a buyer is matched with a seller equals

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

We assume that λ is low enough so that π^{sell} and π^{buy} are below 100%.

Importantly, we assume $\theta > 0$ so 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:

$$\pi_{\rm ho}^s = \lambda \left(\frac{B}{S}\right)^{\frac{1+\theta}{2}} S^{\theta} \text{ and } \pi_{\rm ho}^b = \lambda \left(\frac{S}{B}\right)^{\frac{1-\theta}{2}} S^{\theta}.$$
 (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. We make the standard assumption that $\theta < 1$ to ensure that the "law of diminishing returns" holds for the matching function Eq. (3).¹⁹

The second type of friction arises because it is costly for buyers to participate in asset trading (Vayanos and Wang, 2013). When assets differ in value, participation costs naturally arise because, before making a bid, a buyer must conduct costly valuation analysis, which may involve gathering 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 each trading venue. Intuitively, as *c* increases, a buyer is less likely to participate in multiple trading venues. For simplicity, we assume that *c* is sufficiently high such that each buyer participates in 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 the trading gain, we

¹⁹Specifically, $\theta \ge 1$ implies $\frac{\partial^2 V}{\partial m_b^2} \ge 0$ and $\frac{\partial^2 V}{\partial m_s^2} \ge 0$, indicating that marginal returns increase. In this case, each buyer's trading probability π^{buy} , for instance, weakly *increases* with buyer mass m_b , as shown in Eq. (5), which is unrealistic.

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. 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

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

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 reflects the buyers' bargaining power against the sellers.

3.2 Equilibrium

Next, we describe the traders' choices and the equilibrium.

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

$$\mathcal{M}_{as} = \mathcal{A} = \{1, 2, \cdots, S\}, \qquad (8)$$

where the mass of sellers $s_j = 1$ on each AS submarket $j \in \mathcal{M}_{as}$. 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 .

A seller on any AS submarket trades if she is matched with a buyer. Eq. (7) then implies that a seller on AS submarket j expects to earn

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

where

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

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

Each buyer maximizes expected profit by choosing a trading venues from the set

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

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 is the probability that this buyer trades on submarket *j*. As Eq. (5) implies, the buying probability is

$$\pi_{j}^{b} = \begin{cases} \lambda \left(\frac{1}{b_{j}}\right)^{\frac{1-\theta}{2}} & \forall j \in \mathcal{M}_{as}, \\ 0 & j = 0. \end{cases}$$
(13)

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

As mentioned above, we consider mainly the case in which each buyer participates in at most one AS submarket. To ensure this, we impose the condition $c > 0.25\rho\delta$.²⁰ Moreover, we assume $c < \rho\delta$ to rule out a trivial no-trade equilibrium. If $c \ge \rho\delta$, a participating buyer would always expect to lose $(\pi_j^b \rho \delta - c < 0)$ because the probability of buying, π_j^b , is less than 1. Thus, 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 equilibrium if a vector of buyer masses, $\{b_j : j \in \mathcal{C}^b\}$, satisfies $\sum_{j \in \mathcal{C}^b} b_j = B$ and

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

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

²⁰We show in Lemma A.1 of the appendix that each buyer will choose to participate in at most one venue if the participation cost exceeds $\frac{\rho\delta}{4}$ for every venue. This holds even if the costs vary across venues.

This is a competitive equilibrium in the sense that each buyer takes the equilibrium buyer masses $\{b_j : j \in \mathcal{C}^b\}$ as given. In equilibrium, each buyer weakly prefers her chosen venue to all other options in \mathcal{C}^b .

The equilibrium is as follows:

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

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

where

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

In equilibrium, the AS-only market is fragmented into *S* submarkets of equal size. Each submarket attracts sellers of a distinct asset, each with mass 1, and a disjoint set of buyers, each with mass b^* . To understand why buyers distribute evenly across AS submarkets, note that a buyer expects to earn the same profit $\rho\delta$ from any successful trade, and in equilibrium, each buyer must earn the same expected profit. Therefore, in equilibrium, a buyer on any submarket must trade at the same probability, which implies, based on Eq. (13), that each submarket attracts the same mass of buyers.

Moreover, as Eq. (16) shows, the mass of buyers on any AS submarket cannot exceed \bar{b} ; otherwise, buyers would earn negative expected profits. Therefore, the total mass of buyers participating in all AS submarkets satisfies

$$Sb^* \le S\bar{b}.$$
 (17)

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

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

Thus, if the total mass of buyers *B* exceeds $S\bar{b}$, then $b_0 > 0$, meaning some buyers choose



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

not to participate in trading, and *all* buyers earn zero profit. If $B \le S\overline{b}$, then every buyer participates in trading and earns weakly positive profits.

We illustrate the equilibrium for a market with a total mass of assets S = 4 and a total mass of buyers B = 10 in Fig. 1. In this case, the market fragments into 4 AS submarkets, each attracting buyers with a mass of $b^* = 2$. The mass of non-participating buyers is $b_0 = 2$, as buyers 9 and 10 choose not to participate in any venue and earn zero profits. Since each buyer earns the same profit in equilibrium, every buyer in this example earns 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 successfully sells her asset. To capture the liquidity level of *the whole market*, we use the average probability that an asset is sold across all venues (AS submarkets). Given that 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 commonly used metrics for measuring 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.,* $\pi_i^s = \pi^s$ *, where*

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

the average probability that an asset is sold also equals

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

and the total trading volume equals

$$V = S \cdot \pi^{\text{avg}} = \lambda \left(b^* \right)^{\frac{1+\theta}{2}} \cdot S.$$
(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 . This implies, based on Eq. (9), that every seller earns the same expected profit, i.e., $\psi_j^s = \psi^s$ for any $j \in \mathcal{A}$, 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

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

The loss in liquidity stems from two effects. The first inequality highlights the impact of the cost *c* buyers incur to participate in asset trading. With heterogeneous assets, buyers face participation costs because they must analyze assets before bidding. When expected trading probabilities are low, buyers may opt out, limiting the buyer mass on each submarket to $\bar{b} = (\lambda \rho \delta / c)^{\frac{2}{1-\theta}}$, which reduces liquidity. In contrast, with homogeneous assets, buyers know asset values without incurring analysis costs (*c* = 0). As a result, the cap on buyer mass $\bar{b} \to \infty$ does not bind, and all buyers always participate. The second inequality in Eq. (23) reflects the liquidity externality of pooling traders. Specifically, only 1 unit of an asset is traded on each AS submarket when assets are heterogeneous, whereas all *S* units are traded in a single market with homogeneous assets, i.e., $1 = 1^{\theta} \leq S^{\theta}$. Thus, asset heterogeneity limits the positive network externalities that

enhance liquidity.

Moreover, Theorem 1 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 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 first describe the model setup for the parallel-trading market, where both AS trading and QC trading are allowed. We then derive the parallel-trading equilibrium and compare it to the AS-only market equilibrium discussed in Section 3.

4.1 Setup

As noted in Section 2.1, a TBA-like QC contract does not specify a single deliverable asset but allows delivery from a set of assets. Given our focus on asset value heterogeneity *in value*, we assume a QC contract defines a threshold \underline{v} , making assets QC-eligible if and only if their value satisfies $v \ge \underline{v}$. To account for both QC-eligible and QC-ineligible assets, we assume $\underline{v} > v_{\min}$, ensuring some assets are excluded from QC trading.²¹ In this section, we take v as given.²²

Since both AS and QC contracts are available, sellers of QC-eligible assets, along with all buyers, can choose between the QC contract and AS contracts. Consequently, the set of assets, \mathcal{A} , is endogenously partitioned into the set of assets traded on the QC market,

²¹Some MBS, for instance, are ineligible for TBA trading due to factors such as containing a high proportion of jumbo loans that exceed conforming size limits.

²²In Internet Appendix IA.1, we analyze the market design considerations involved in selecting \underline{v} .

 $\widetilde{\mathscr{M}}_{qc}$, and the set of assets traded on AS submarkets, $\widetilde{\mathscr{M}}_{as}.^{23}$ That is,

$$\widetilde{\mathcal{M}}_{as} = \mathscr{A} - \widetilde{\mathcal{M}}_{qc}.$$
(26)

It follows that

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

where $S_{qc} = |\widetilde{\mathcal{M}}_{as}|$ and $S_{as} = |\widetilde{\mathcal{M}}_{as}|$ represent the masses of assets on the QC market and on all AS submarkets, respectively. Additionally, 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, $B_{as} = \sum_{j \in \widetilde{\mathcal{M}}_{as}} \widetilde{b}_j$ and $\widetilde{b}_0 = B - B_{as} - B_{qc}$.

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

$$P_{\rm qc} = \begin{cases} \nu_{\rm qc} + \delta & \text{with probability } 1 - \rho, \\ \nu_{\rm qc} & \text{with probability } \rho, \end{cases}$$
(28)

where

$$\nu_{\rm qc} := \mathbf{E} \left[\nu_{\tau} \left| \tau \in \widetilde{\mathcal{M}}_{\rm qc} \right. \right] \tag{29}$$

denotes 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 on the QC market is matched, she expects to earn $\rho\delta$, which is the same profit she would earn in an AS trade (see Eq. (7) and the related discussions).

We assume that trading in the parallel-trading market, both on AS and QC venues,

²³We use tilde to indicate that a variable represents a characteristic of the market in which both AS and QC trading are available.

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

$$\widetilde{\pi}_{\rm qc}^b = \lambda \left(\frac{S_{\rm qc}}{B_{\rm qc}}\right)^{\frac{1-\theta}{2}} S_{\rm qc}^\theta \quad \text{and} \quad \widetilde{\pi}_{\rm qc}^s = \lambda \left(\frac{B_{\rm qc}}{S_{\rm qc}}\right)^{\frac{1-\theta}{2}} B_{\rm qc}^\theta; \tag{30}$$

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

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

As before, the trading probability for non-participating buyers $\widetilde{\pi}^b_0$ equals 0.

Moreover, we assume that for a buyer to participate in QC trading, she must pay c, the same cost required for participating in the AS trading of an asset.²⁴

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 a 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

$$\widetilde{\psi}_{\mathrm{as},j}^{s} = \widetilde{\pi}_{\mathrm{as},j}^{s} \mathbf{E} \left[P_{\mathrm{as}}(\nu_{j}) - \nu_{j} \right] = \widetilde{\pi}_{\mathrm{as},j}^{s} (1 - \rho) \delta.$$
(32)

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

$$\widetilde{\mathscr{C}}_{j}^{s} := \left\{ qc, j \right\}, \tag{33}$$

²⁴In Appendix B.1, we extend the model and show that our results remain qualitatively valid if the cost to participate in QC trading c_{qc} does not exceed cS_{qc}^{θ} . In practice, for buyers, participating in the QC market is likely not significantly more costly than participating in an AS submarket. For instance, in the MBS market, while TBA buyers need to estimate the average value of TBA MBS, they can rely on quotes from multiple dealers for the same contract, which reduces their analysis costs. In contrast, quotes from a single seller for an SP MBS are asset-specific and provide limited information for pricing another SP MBS sold by a different seller, requiring buyers to analyze each asset they trade.

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

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

where $\tilde{\pi}_{qc}^s$ and $\tilde{\pi}_{as,j}^s$ 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 offers higher liquidity, their assets could fetch higher prices on AS submarkets. In equilibrium, sellers of high-value assets, whose values exceed the endogenous threshold \bar{v} , opt for AS trading despite its lower liquidity, because their assets are significantly more valuable than P_{qc} .

Each buyer's choice set is

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

where 0 represents non-participation, "qc" represents the QC market, and $\widetilde{\mathcal{M}}_{as}$ 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

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

where $\tilde{\pi}_{qc}^{b}$ and $\tilde{\pi}_{as,j}^{b}$ are defined in Eqs. (30) and (31). The pricing functions we assume in Eqs. (7) and (28) then imply that a buyer's expect profit

$$\widetilde{\psi}_{k}^{b} = \widetilde{\pi}_{k}^{b} \rho \delta - c \cdot \mathbb{1}_{\{k \neq 0\}}, \qquad (37)$$

which depends on her trading venue k through trading probability $\tilde{\pi}_{k}^{b}$.

To derive the equilibrium, we make a tie-breaking assumption that sellers prefer the QC market over an AS submarket when they are indifferent between the two.²⁵ We then state the equilibrium conditions as follows:

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 ≥ <u>v</u>), her trading venue choice ε_j maximizes her expected profit:

$$\widetilde{\psi}(\varepsilon_j, v_j) \ge \max_{e \in \widetilde{\mathscr{C}}_j^s, e \neq \varepsilon_j} \quad \widetilde{\psi}(e, v_j)$$
(38)

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

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

$$\widetilde{\psi}_{k}^{b} \ge \max_{k' \in \widetilde{\mathscr{C}}^{b}, k' \neq k} \quad \widetilde{\psi}_{k'}^{b}, \tag{39}$$

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

As in the benchmark model, this is a competitive equilibrium where every trader takes the equilibrium buyer masses $b_k : k \in \mathcal{C}^b$ as given. The key difference here is that each buyer now has one more venue to choose from—the QC market—alongside the AS submarkets and the option of non-participation. Additionally, each seller, taking the venue choices of other sellers as given, chooses her trading venue. Furthermore, the trading probabilities for both buyers $(\tilde{\pi}_k^b)$ and sellers $(\tilde{\pi}_{qc}^s \text{ and } \tilde{\pi}_{as,j}^s)$ depend on the equilibrium set of sellers who choose the QC market and the equilibrium masses of buyers on each trading venue.

²⁵This assumption implies that all units of an asset are traded either entirely through AS trading or entirely through QC trading. Hence, at least one asset is traded via the QC market (i.e., $S_{qc} \ge 1$).

For ease of presenting the equilibrium, we define two increasing and non-negative functions:

$$\mu(x) := x \left(x^{\frac{2\theta}{1-\theta}} - 1 \right) \qquad \text{for } x \ge 1 \tag{40}$$

$$\eta(v^*) := \mathbf{E}[v|v \in [\underline{v}, v^*]] + (1-\rho)\delta\left(1 - \frac{1}{\left(S \cdot \Pr\left\{v \in [\underline{v}, v^*]\right\}\right)^{\frac{2\theta}{1-\theta}}}\right) \quad \text{for } v^* > \underline{v}.$$
(41)

The equilibrium is as follows.

Theorem 2 (**Parallel-trading equilibrium**). *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 $v_j \in [\underline{v}, \overline{v}]$ and the AS submarket for asset j if $v_j \in [v_{\min}, \underline{v}) \cup (\overline{v}, v_{\max}]$, where \overline{v} satisfies $\overline{v} = \min \{\eta(\overline{v}), v_{\max}\}$. That is,

$$\widetilde{\mathcal{M}}_{qc} = \left\{ j : j \in \mathscr{A} \text{ and } \underline{v} \le v_j \le \overline{v} \right\},\tag{42}$$

$$\widetilde{\mathcal{M}}_{as} = \mathscr{A} - \widetilde{\mathcal{M}}_{qc}.$$
(43)

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

$$B_{\rm qc} = (Sq)^{\frac{1+\theta}{1-\theta}} \cdot \tilde{b}^*, \tag{44}$$

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

$$\widetilde{b}_j = \widetilde{b}^* \quad \forall j \in \widetilde{\mathcal{M}}_{\mathrm{as}},\tag{45}$$

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

$$\widetilde{b}_0 = \max\left\{B - \left(S + \mu(Sq)\right)\overline{b}, 0\right\},\tag{46}$$

where

$$q := \Pr\left\{v \in [\underline{v}, \bar{v}]\right\},\tag{47}$$

$$\widetilde{b}^* := \min\left\{\frac{B}{S+\mu(Sq)}, \overline{b}\right\},\tag{48}$$

and \overline{b} is defined in Eq. (15).

A key endogenous parameter, \bar{v} , which represents the upper bound for the values of assets traded on the QC market, characterizes the equilibrium.²⁶ A seller chooses the QC market only if the value of her asset lies in the interval $[\underline{v}, \bar{v}]$. Sellers of low-value assets, whose values are below \underline{v} , must 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 determine the fraction of assets sold on the QC market, denoted by q. Then, the mass of sellers on the QC market, and the total mass of sellers on all AS submarkets equal, respectively,

$$S_{qc} = Sq$$
 and $S_{as} = S(1-q)$. (49)

In addition, because each AS submarket attracts buyers of mass \tilde{b}^* , the total mass of buyers across all AS submarkets equals

$$B_{\rm as} = S_{\rm as}\tilde{b}^*. \tag{50}$$

Fig. 2 illustrates an example parallel-trading equilibrium in which q = 0.5, S = 4, $B_{qc} = 7$, $\tilde{b}^* = 1$, and $\tilde{b}_0 = 1$. 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, as illustrated in Fig. 1, the buyer-to-seller ratio in the AS-only market (which equals 2) lies between them. In addition, more buyers participate in trading in

²⁶While multiple equilibria may exist, our results hold for any equilibrium.



Figure 2. Illustration of the Parallel-Trading Equilibrium

the parallel trading market than in the AS-only market (9 > 8). 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

$$\hat{b}^* \le b^* \le B_{\rm qc} / S_{\rm qc},\tag{51}$$

where \tilde{b}^* and B_{qc}/S_{qc} represent the buyer-to-seller ratios on each AS submarket and on the QC market in the parallel-trading market equilibrium, respectively, and b^* is 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.

$$B - b_0 \ge B - b_0. \tag{52}$$

To understand why, in the parallel-trading equilibrium, the buyer-to-seller ratio on the QC market, B_{qc}/S_{qc} , is weakly greater than the ratio on each AS submarket, \tilde{b}^* , consider the hypothetical situation where $B_{qc}/S_{qc} < \tilde{b}^*$. In this scenario, the probability that a buyer on the QC market trades, $\tilde{\pi}_{qc}^b$ (given in Eq. (30)), would be strictly higher than the probability that a buyer on an AS submarket trades, $\tilde{\pi}_{as,j}^b$ (given in Eq. (31)), due to both the strictly lower buyer-to-seller ratio ($B_{qc}/S_{qc} < \tilde{b}^*$) and the liquidity externality from pooling multiple types of assets (i.e., $S_{qc}^{\theta} \ge 1$). As a result, a buyer on the QC market would earn strictly higher profit than a buyer on an AS submarket, as shown in Eq. (37). This cannot be an equilibrium because buyers can choose venues freely. Furthermore, Corollary 2 shows that, in a parallel-trading market, the buyer-to-seller ratio is lower on AS submarkets but higher on the QC market, relative to the buyer-to-seller ratio in the AS-only market (b^*).

In addition, fewer buyers choose not to participate ($\tilde{b}_0 \leq b_0$), which can be derived by comparing Eqs. (18) and (46) given that $S + \mu(Sq) \geq S \geq 1$. Intuitively, more buyers participate because each participating buyer is more likely to trade after paying the same participation cost and earns weakly more profit.

4.3 Market Liquidity

In this section, we first derive liquidity levels in the parallel-trading equilibrium and then compare them with liquidity levels in the AS-only equilibrium.

Liquidity levels in the parallel-trading equilibrium. As in the benchmark model, we define the liquidity level of a trading venue as the probability that an asset on that venue is sold, and the overall market liquidity as the average probability that assets across all venues are sold. With a fixed set of assets, this average probability is proportional to the expected total trading volume and turnover across all venues. Using Theorem 2, we derive the liquidity levels in the parallel-trading equilibrium 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: $\tilde{\pi}_{as,i}^s = \tilde{\pi}_{as}^s$ where

$$\widetilde{\pi}_{as}^{s} = \lambda (\widetilde{b}^{*})^{\frac{1+\theta}{2}};$$
(53)

• an asset on the QC market is sold with probability

$$\widetilde{\pi}_{qc}^{s} = \lambda(\widetilde{b}^{*})^{\frac{1+\theta}{2}}(Sq)^{\frac{2\theta}{1-\theta}};$$
(54)

• the average selling probability of all assets equals

$$\widetilde{\pi}^{\text{avg}} = \lambda(\widetilde{b}^*)^{\frac{1+\theta}{2}} \frac{S + \mu(Sq)}{S},\tag{55}$$

and the total expected trading volume across all venues equals

$$\widetilde{V} = S\widetilde{\pi}^{\text{avg}} = \lambda(\widetilde{b}^*)^{\frac{1+\theta}{2}}(S + \mu(Sq)).$$
(56)

We derive Eq. (53) by substituting Eq. (45) into Eq. (31), and Eq. (54) by substituting Eqs. (44) and (49) into Eq. (30). As Eq. (53) shows, in the parallel-trading equilibrium, AS submarkets for both QC-eligible and QC-ineligible assets remain equally liquid. Furthermore, comparing Eq. (53) and Eq. (54), we find $\tilde{\pi}_{as}^s \leq \tilde{\pi}_{qc}^s$ (since $Sq \geq 1$), which is consistent with empirical findings from the agency MBS market in Gao et al. (2017).

Comparison with the AS-only market. Next, we compare the liquidity levels in the parallel-trading equilibrium, as presented in Theorem 2, with those in the AS-only equilibrium, as outlined in Theorem 1.

We begin by comparing the liquidity levels for individual assets.

Corollary 4 (**Comparison of liquidity of individual assets**). *Compared with the AS-only market, in the parallel-trading equilibrium, assets traded on the QC market are more liq-uid, while assets traded on AS submarkets are less liquid. Formally, we have*

$$\widetilde{\pi}_{as}^{s} \le \pi^{s} \le \widetilde{\pi}_{qc}^{s}.$$
(57)

We derive the first inequality in Eq. (57) by comparing Eq. (53) with Eq. (19), noting from Corollary 2 that $\tilde{b}^* \leq b^*$. In the parallel-trading equilibrium, the buyer-to-seller ratios on AS submarkets are lower compared to the AS-only equilibrium, leading to reduced liquidity of assets on these submarkets. Thus, the QC market effectively siphons liquidity away from the AS submarkets.

We derive the second inequality in Eq. (57) as follows:

$$\widetilde{\pi}_{\rm qc}^{s} = \lambda \left(\frac{B_{\rm qc}}{S_{\rm qc}}\right)^{\frac{1-\theta}{2}} \cdot B_{\rm qc}^{\theta} \ge \lambda \left(b^*\right)^{\frac{1-\theta}{2}} \left(\frac{B_{\rm qc}}{S_{\rm qc}} \cdot S_{\rm qc}\right)^{\theta} \ge \lambda \left(b^*\right)^{\frac{1-\theta}{2}} \left(b^* \cdot 1\right)^{\theta} = \pi^s.$$
(58)

We derive the first inequality in Eq. (58) using the increased buyer-to-seller ratio $(B_{qc}/S_{qc} \ge b^*)$, and the second inequality using both the increased buyer-to-seller ratio $(B_{qc}/S_{qc} \ge b^*)$ and the liquidity benefit of pooling more traders $(S_{qc}^{\theta} \ge 1)$.

We then compare the overall liquidity of the parallel-trading equilibrium and that of the AS-only equilibrium. To do this, we introduce an auxiliary result concerning buyers. Specifically, plugging Eqs. (44), (45) and (49) into Eq. (31), we find that in equilibrium a buyer on any venue trades with the same probability, i.e., $\tilde{\pi}_{qc}^b = \tilde{\pi}_{as,j}^b = \tilde{\pi}^b$ for any $j \in \widetilde{\mathcal{M}}_{as}$, where

$$\widetilde{\pi}^{b} = \lambda \left(\frac{1}{\widetilde{b}^{*}}\right)^{\frac{1-\theta}{2}}.$$
(59)

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 (59) and Corollary 2, that

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

so buyers on all venues are more likely to trade in the parallel-trading equilibrium than in the AS-only equilibrium.

Lemma 1 (Buyers' trading probability). $\tilde{\pi}^b \ge \pi^b$.

Building on this auxiliary result, we derive the following result.

Corollary 5 (**Comparison of liquidity of the whole market**). *The overall liquidity of all assets is greater in the parallel-trading equilibrium than in the AS-only equilibrium. For-mally, we have*

$$\widetilde{V} \ge V \quad and \quad \widetilde{\pi}^{avg} \ge \pi^{avg}.$$
 (61)

The total trading volume is higher in the parallel-trading equilibrium compared to the AS-only equilibrium, i.e., $\tilde{V} \ge V$. This is because, as Corollary 2 and Lemma 1 imply, more buyers participate in trading in the parallel-trading equilibrium $(B - \tilde{b}_0 \ge B - b_0)$, and each participating buyer is more likely to trade $(\tilde{\pi}^b \ge \pi^b)$. The total expected trading volume is the product of the probability that each participating buyer trades and the mass of all participating buyers. Therefore, $\tilde{V} = (B - \tilde{b}_0)\tilde{\pi}^b \ge (B - b_0)\pi^b = V$.

Additionally, since 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*, is also higher in the parallel-trading equilibrium. Specifically, $\tilde{\pi}^{avg} = \tilde{V}/S \ge V/S = \pi^{avg}$.

In summary, compared to the AS-only equilibrium, assets sold on the QC market are more liquid in the parallel-trading equilibrium, while assets sold on AS submarkets are less liquid. Overall, the liquidity improvement for QC assets more than compensates for the reduced liquidity of AS assets, resulting in a higher average liquidity for all assets in the parallel-trading equilibrium than in the AS-only equilibrium.

4.4 Trader Profits and Social Welfare

In this section, we compare traders' profits and social welfare in the parallel-trading equilibrium with those in the AS-only equilibrium.

First, we compare buyers' profits. According to Eqs. (25) and (37), a buyer in the ASonly equilibrium earns a profit of $\psi^b = \pi^b \rho \delta - c$, whereas a buyer in the parallel-trading equilibrium earns $\tilde{\psi}^b = \tilde{\pi}^b \rho \delta - c$. Since $\tilde{\pi}^b \ge \pi^b$ according to Lemma 1, we obtain the following result.

Corollary 6 (Comparison of buyers' profits). A buyer earns more profits in the paralleltrading equilibrium compared to a buyer in the AS-only equilibrium, i.e., $\tilde{\psi}^b \ge \psi^b$.

Second, we compare sellers' profits. According to Eqs. (22), (32) and (34), each seller in the AS-only market earns a profit of $\psi^s = \pi^s (1 - \rho)\delta$, whereas a seller on any AS submarket in the parallel-trading equilibrium earns a profit of

$$\widetilde{\psi}_{as}^{s} = \widetilde{\pi}_{as}^{s} (1 - \rho) \delta.$$
(62)

Because $\tilde{\pi}_{as}^s \leq \pi^s$ according to Corollary 4, we have $\tilde{\psi}_{as}^s \leq \psi^s$, meaning that AS sellers in the parallel-trading equilibrium are less likely to trade and, as a result, earn lower profits compared to AS sellers in the AS-only equilibrium.

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

$$\widetilde{\psi}^{s}(qc, \nu_{j}) = \widetilde{\pi}_{qc}^{s} \left(\mathbf{E}[P_{qc}] - \nu_{j} \right) = \widetilde{\pi}_{qc}^{s} \left((1 - \rho)\delta + \nu_{qc} - \nu_{j} \right), \tag{63}$$

which decreases with the value of her asset v_j because every asset on the QC market is sold at the same price P_{qc} .

To compare the profits of sellers choosing QC trading in the parallel-trading equilibrium with those of sellers in the AS-only equilibrium, we solve $\tilde{\psi}^{s}(qc, v) = \psi^{s}$ and find that

$$v^* = v_{\rm qc} + (1 - \rho)\delta\left(1 - \frac{\pi^s}{\tilde{\pi}_{\rm qc}^s}\right). \tag{64}$$

Thus, if a QC seller's asset is worth v^* , she earns the same profit as in the AS-only market. Because $\tilde{\psi}^s(qc, v_j)$ decreases with v_j , a QC seller in the parallel-trading equilibrium earns more profit than a seller in the AS-only equilibrium ($\tilde{\psi}^s(qc, v_j) > \psi^s$) if her asset is less valuable than v^* . Conversely, she earns less profit ($\tilde{\psi}^s(qc, v_j) < \psi^s$) if her asset is more valuable than v^* .

Note that sellers with asset values in the interval $(v^*, \bar{v}]$ choose QC trading in the parallel-trading equilibrium, but they earn lower profits than they would in the AS-only equilibrium. To understand this result, recall from Corollary 4 that AS submarkets in the parallel-trading equilibrium are less liquid than AS submarkets in the AS-only equilibrium. As a result, the outside option of selling on AS submarkets becomes less valuable for QC sellers in the parallel-trading equilibrium. This negative effect outweighs the liquidity benefits of trading on the QC market for sellers with assets in the interval $(v^*, \bar{v}]$, leading to lower profits compared to what they would earn in the AS-only equilibrium.

It is worth noting that it is possible for *all* QC sellers in the parallel-trading equilibrium to earn higher profits than they would in the AS-only equilibrium. Specifically, if *all* QC-eligible assets are sold through QC trading ($\bar{v} = v_{max}$), then *all* sellers who opt for QC trading will earn higher profits than they would in the AS-only equilibrium. However, 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 $v^* \leq \bar{v}$. In this case, QC sellers with asset values in the interval [v^* , \bar{v}] will earn weakly lower profits than they could have earned in the AS-only equilibrium.

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

Corollary 7 (Comparison of seller's profits). Compared to the AS-only market equilib-



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

rium, in the parallel-trading market equilibrium, a seller on AS markets earns less profit $(\tilde{\psi}_{as}^s \leq \psi_{as}^s)$, while a seller on the QC market earns more profit if her asset's value $v_j < v^*$ and earns 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

$$\Omega = \underbrace{(B - b_0)(\rho \delta \pi^b - c)}_{\text{all buyers' total profit}} + \underbrace{S\pi^{\text{avg}}(1 - \rho)\delta}_{\text{all sellers' total profit}},$$
(65)

where $B - b_0$ equals the total mass of participating buyers, $\rho \delta \pi^b - c$ equals the profit of each participating buyer (see 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

$$\widetilde{\Omega} = \underbrace{(B - \widetilde{b}_{0})(\rho \delta \widetilde{\pi}^{b} - c)}_{\text{all buyers' total profit}} + \underbrace{S(1 - q)\widetilde{\pi}_{as}^{s}(1 - \rho)\delta}_{\text{AS sellers' total profit}} + \underbrace{Sq\widetilde{\pi}_{qc}^{s} \mathbf{E}\left[(1 - \rho)\delta + v_{qc} - v_{j}|v_{j} \in [\underline{v}, \overline{v}]\right]}_{\text{QC sellers' total profit}}$$

$$= \underbrace{(B - \widetilde{b}_{0})(\rho \delta \widetilde{\pi}^{b} - c)}_{\text{all buyers' total profit}} + \underbrace{S(1 - q)\widetilde{\pi}_{as}^{s}(1 - \rho)\delta}_{\text{AS sellers' total profit}} + \underbrace{Sq\widetilde{\pi}_{qc}^{s}(1 - \rho)\delta}_{\text{QC sellers' total profit}}$$

$$= \underbrace{(B - \widetilde{b}_{0})(\rho \delta \widetilde{\pi}^{b} - c)}_{\text{all buyers' total profit}} + \underbrace{S\widetilde{\pi}^{avg}(1 - \rho)\delta}_{\text{all sellers' total profit}}$$
(66)

where $B - \tilde{b}_0$ equals the total mass of participating buyers, $\rho \delta \tilde{\pi}^b - c$ equals the profit of each participating buyer, S(1 - q) equals the total mass of sellers on AS submarkets, $\tilde{\pi}_{as}^s(1-\rho)\delta$ equals the expected profit of each seller on AS submarkets (as Eq. (62) shows), Sq equals the total mass of sellers on the QC submarket, $\tilde{\pi}_{qc}^s \mathbb{E}[(1-\rho)\delta + v_{qc} - v_j | v_j \in [\underline{v}, \bar{v}]]$ equals the average expected profit of QC sellers (note that the expectation is taken for $v_j \in [\underline{v}, \bar{v}]$), and $\tilde{\pi}^{avg}$ equals the average asset selling probability based on Eq. (55). Because assets on the QC market are priced uniformly according to their average value v_{qc} , we have that $\mathbb{E}[(1-\rho)\delta + v_{qc} - v_j | v_j \in [\underline{v}, \bar{v}]] = (1-\rho)\delta$, so the average expected profit of QC sellers equals $\tilde{\pi}_{qc}^s(1-\rho)\delta$.

Comparing Eqs. (65) and (66), we observe that $\tilde{\Omega} \ge \Omega$ for two reasons. First, as Lemma 1 shows, in the parallel-trading equilibrium, more buyers participate $(B - \tilde{b}_0 \ge B - b_0)$, and each is more likely to trade $(\tilde{\pi}^b \ge \pi^b)$, resulting in higher total profits for buyers. Second, as Corollary 5 shows, in the parallel-trading equilibrium, assets are more liquid on average $(\tilde{\pi}^{avg} \ge \pi^{avg})$, leading to higher total profits for sellers. Thus, we obtain the following result:

Corollary 8 (Comparison of Social Welfare). Social welfare is higher in the paralleltrading equilibrium than in the AS-only equilibrium ($\tilde{\Omega} \ge \Omega$).

5 Conclusion

We develop a model of over-the-counter (OTC) markets involving assets with heterogeneous fundamental values. First, we show that asset heterogeneity reduces liquidity by limiting the network externalities from pooling traders together. Motivated by the markets for agency MBS, we then analyze a parallel-trading market that allows both standard asset-specific (AS) trading and a TBA-like quasi-consolidated (QC) trading mechanism, where heterogeneous assets within a cohort are sold at a uniform price. We find that, compared to the AS-only market, the parallel-trading market has higher total trading volume and greater social welfare, but lower liquidity for assets traded through standard asset-specific contracts.

Asset heterogeneity is a key and common feature of OTC markets. Our theoretical framework provides a foundation for analyzing liquidity and trading design in markets with heterogeneous assets, such as corporate bonds, municipal bonds, MBS, and assetbacked securities. For instance, one may use our framework to analyze key market design issues, such as whether to implement QC trading and how to define its specifications.²⁷ Future research in this direction will be crucial for deepening our understanding of OTC markets.

Appendix

A Proofs for Results in the Main Text

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 \mathcal{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 $\mathscr{I} = \{1, \dots, I\}$ and $\pi_1^b \le \pi_i^b$ for $i \in \mathscr{I}$. The buyer expects to earn

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

²⁷For illustration, we analyze the choice of QC-eligibility requirements in Section Appendix IA.1, drawing on the Securities Industry and Financial Markets Association (SIFMA) restrictions on TBA-eligibility for high-balance loans in 2008 (Vickery and Wright, 2011).

If the buyer quits trading venue 1, she would earn

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

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

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

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 $B \ge S\bar{b}$, 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 $\forall j \in \mathcal{A}$. 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 $\tilde{b}^* = B_{\rm as}/S_{\rm as} = B_{\rm qc}/S_{\rm qc}^{\frac{1+\theta}{1-\theta}}$ and $\tilde{\pi}_{\rm qc}^s = \tilde{\pi}_{\rm as}^s S_{\rm qc}^{\frac{2\theta}{1-\theta}}$.

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

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

and

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

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

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

Proof of Lemma A.4. Eq. (34) implies that $\tilde{\psi}(qc, v_j) - \tilde{\psi}(j, v_j) = \tilde{\pi}_{qc}^s(v_{qc} - v_j) + (\tilde{\pi}_{qc}^s - \tilde{\pi}_{as,j}^s)(1-\rho)\delta$, which is non-negative if and only if $v_j \le v_{qc} + \left(1 - \frac{\tilde{\pi}_{as,j}^s}{\tilde{\pi}_{qc}^s}\right)(1-\rho)\delta$.

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 $\tilde{b} = B_{\rm as}/S_{\rm as} = B_{\rm qc}/S_{\rm qc}^{\frac{1+\theta}{1-\theta}}$. By definition, $S_{\rm as} = (1-q)S$ and $S_{\rm qc} = qS$. Thus, the total mass of participating buyers

$$B - \widetilde{b}_0 = B_{\rm as} + B_{\rm qc} = \widetilde{b}\left(S_{\rm as} + S_{\rm qc}^{\frac{1+\theta}{1-\theta}}\right) = \widetilde{b}\left(S + \mu(Sq)\right) \le B,\tag{A.6}$$

which implies that $\tilde{b} \leq \frac{B}{S+\mu(Sq)}$. In addition, Lemma A.2 shows that $\tilde{b} \leq \tilde{b}$. Thus, $\tilde{b} \leq \tilde{b}^* = \min\left\{\frac{B}{S+\mu(Sq)}, \tilde{b}\right\}$. If $\tilde{b} < \tilde{b}$, then $\tilde{\pi}^b_{as}\rho\delta - c > 0$ and a buyer on an AS submarket earns strictly positive profit, so all buyers participate ($\tilde{b}_0 = 0$), which implies that $\tilde{b} = \frac{B}{S+\mu(Sq)}$; if $\tilde{b} < \frac{B}{S+\mu(Sq)}$, then some buyers do not participate ($\tilde{b}_0 = B - \tilde{b}(S+\mu(Sq)) > 0$), so all buyers earn zero profit and $\tilde{b} = \tilde{b}$. Thus, $\tilde{b} = \tilde{b}^*$.

Second, we find sellers' venue choices. Sellers of QC-ineligible assets can use only AS trading. Consider a seller of a QC-eligible asset whose value $v' \ge \underline{v}$. Suppose that other buyers and sellers follow the equilibrium strategy. If $v' \in [\underline{v}, \overline{v}]$, the seller chooses QC trading because no buyers participate in the AS submarket of this asset. If $v' > \overline{v}$, then, $\overline{v} < v' \le v_{\text{max}}$, which, given the equilibrium condition that $\overline{v} = \min \{\eta(\overline{v}), v_{\text{max}}\}$, implies that $v_{\text{max}} > \overline{v} = \eta(\overline{v})$ and that $v' > \eta(\overline{v})$. Additionally,

$$\eta(\bar{\nu}) = \nu_{\rm qc} + (1-\rho)\delta\left(1 - \frac{1}{S_{\rm qc}^{\frac{2\theta}{1-\theta}}}\right) = \nu_{\rm qc} + \left(1 - \frac{\tilde{\pi}_{\rm as,j}^s}{\tilde{\pi}_{\rm qc}^s}\right)(1-\rho)\delta. \tag{A.7}$$

Thus, based on Lemmas A.3 and A.4, this seller prefers the AS market. Therefore, these seller also follow the equilibrium strategy in choosing trading venue. \Box

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

Proof of Lemma A.5. By definition, $q \le 1$ and $\mu(1) = 0$. Additionally, as Footnote 25 explains, $Sq \ge 1$. 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 Corollaries 2, 4 and 5 and Lemma 1. Because $\mu(Sq) \ge 0$, we have that $\tilde{b}^* \le b^*$ and $\tilde{b}_0 \le b_0$. Because $\tilde{b}^* \le b^*$, we have that $\tilde{\pi}_{as}^s \le \pi^s$ and $\tilde{\pi}^b \ge \pi^b$. Thus, $\tilde{V} = (B - \tilde{b}_0)\tilde{\pi}^b \ge (B - b_0)\pi^b = V$ and $\tilde{\pi}^{avg} = \tilde{V}/S \ge V/S = \pi^{avg}$.

We then show that $b^* \leq B_{qc}/S_{qc}$. Based on Lemma A.5, we have that $S \cdot (Sq)^{\frac{2\theta}{1-\theta}} \geq S + \mu(Sq)$. Thus,

$$\frac{B}{S+\mu(Sq)}(Sq)^{\frac{2\theta}{1-\theta}} = \frac{B}{S}\frac{S}{S+\mu(Sq)}(Sq)^{\frac{2\theta}{1-\theta}} \ge \frac{B}{S}.$$
(A.8)

It follows that

$$\frac{B_{\rm qc}}{S_{\rm qc}} = \tilde{b}^* S_{\rm qc}^{\frac{2\theta}{1-\theta}} = \min\left\{\frac{B}{S+\mu(Sq)}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}\right\} (Sq)^{\frac{2\theta}{1-\theta}} \\
\geq \min\left\{\frac{B}{S+\mu(Sq)}(Sq)^{\frac{2\theta}{1-\theta}}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}\right\} \geq \min\left\{\frac{B}{S}, \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}\right\} = b^*.$$
(A.9)

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

Proof of Corollaries 6 to 8. The proof of Corollary 6 is provided in the main text before the result. Then, we show by contradiction that $v^* \in (\bar{v}, v_{\text{max}}]$ cannot hold. If $v^* \in (\bar{v}, v_{\text{max}}]$, then

$$\nu^* = \nu_{\rm qc} + (1-\rho)\delta\left(1 - \frac{\pi^s}{\widetilde{\pi}_{\rm qc}^s}\right) > \nu_{\rm qc} + (1-\rho)\delta\left(1 - \frac{\widetilde{\pi}_{\rm as}^s}{\widetilde{\pi}_{\rm qc}^s}\right),\tag{A.10}$$

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

B Extensions

In this section, we extend the main model to consider two more general cases: (1) QC trading is more costly to participate in than AS trading, and (2) trading gains differ across assets.

B.1 Higher Cost to Participate in QC Trading

In the main model, buyers incur the same cost to participate in any venue. However, buyers may presumably incur higher analysis costs in the QC market because more assets traded together. Thus, in this section, we allow a higher cost, c_{qc} , for participating in the QC market compared to the AS submarket, i.e., $c_{qc} \ge c$. Additionally, we allow c_{qc} to increase with the number of assets traded on the QC market, S_{qc} .

We show that if $c_{qc} \leq c S_{qc}^{\theta}$, all results concerning the qualitative comparison between the AS-only and parallel-trading equilibria(Corollaries 2, 4 to 6 and 8 and Lemma 1) remain valid. Intuitively, in the extended model, concentrating more assets in the QC market enhances liquidity (captured by S_{qc}^{θ}) but also raises buyer participation costs (captured by c_{qc}/c). When $S_{qc}^{\theta} \geq c_{qc}/c$, the liquidity benefit outweighs the higher costs, leading to disproportionately more buyers than sellers choosing QC trading. This raises the buyer-to-seller ratio on the QC market and lowers it on AS submarkets ($B_{qc}/S_{qc} \geq b^* \geq \tilde{b}^*$), improving liquidity on the QC market and worsening it on AS submarkets $(\tilde{\pi}_{qc}^s \ge \pi^s \ge \tilde{\pi}_{as}^s)$. Other results then follow. In the rest of this section, we prove these results.

Lemma A.1. If in equilibrium the QC market and AS-submarket j both attract some buyers, then

$$\frac{B_{\rm qc}/S_{\rm qc}}{\widetilde{b}_j} = \left(\frac{S_{\rm qc}^\theta}{C_{\rm qa}}\right)^{\frac{2}{1-\theta}}, \quad \frac{\widetilde{\pi}_{\rm qc}^b}{\widetilde{\pi}_{\rm as,j}^b} = C_{\rm qa}, \quad \frac{\widetilde{\pi}_{\rm qc}^s}{\widetilde{\pi}_{\rm as,j}^s} = \left(\frac{S_{\rm qc}^{\frac{2\theta}{1+\theta}}}{C_{\rm qa}}\right)^{\frac{1+\theta}{1-\theta}}, \tag{A.1}$$

where

$$C_{\rm qa} := \frac{c_{\rm qc} + \widetilde{\psi}^b}{c + \widetilde{\psi}^b} \tag{A.2}$$

and $\tilde{\psi}^b$ equals the expected profit of each buyer.

Proof of Lemma A.1. Because every participating buyer earns the same profit in equilibrium, $\tilde{\psi}^b = \tilde{\pi}^b_{qc} \rho \delta - c_{qc} = \tilde{\pi}^b_{as,j} \rho \delta - c$, which implies that $\tilde{\pi}^b_{qc} / \tilde{\pi}^b_{as,j} = C_{qa}$. By definition, $\tilde{\pi}^b_{as,j} = \lambda \left(\frac{1}{\tilde{b}_j}\right)^{\frac{1-\theta}{2}}$, $\tilde{\pi}^b_{qc} = \lambda \left(\frac{S_{qc}}{B_{qc}}\right)^{\frac{1-\theta}{2}} S^{\theta}_{qc}$, $\tilde{\pi}^s_{as,j} = \lambda \left(\tilde{b}_j\right)^{\frac{1+\theta}{2}}$, $\tilde{\pi}^s_{qc} = \lambda \left(\frac{B_{qc}}{S_{qc}}\right)^{\frac{1+\theta}{2}} S^{\theta}_{qc}$. Thus, we have that

$$\frac{B_{\rm qc}/S_{\rm qc}}{\widetilde{b}_j} = \left(\frac{\widetilde{\pi}_j^b}{\widetilde{\pi}_{\rm qc}^b}\right)^{\frac{2}{1-\theta}} S_{\rm qc}^{\frac{2\theta}{1-\theta}} = \left(\frac{S_{\rm qc}^\theta}{C_{\rm qa}}\right)^{\frac{2}{1-\theta}}$$
(A.3)

and

$$\frac{\widetilde{\pi}_{\rm qc}^s}{\widetilde{\pi}_{\rm as,j}^s} = \left(\frac{B_{\rm qc}/S_{\rm qc}}{\widetilde{b}_j}\right)^{\frac{1+\theta}{2}} S_{\rm qc}^{\theta} = \left(\frac{S_{\rm qc}^{\theta}}{C_{\rm qa}}\right)^{\frac{1+\theta}{1-\theta}} S_{\rm qc}^{\theta} = \left(\frac{S_{\rm qc}^{\frac{2\theta}{1+\theta}}}{C_{\rm qa}}\right)^{\frac{1+\theta}{1-\theta}}.$$
(A.4)

Lemma A.2. In equilibrium, buyer mass on any AS submarket j equals

$$\tilde{b}_j = \tilde{b}^*, \tag{A.5}$$

1.0

the buyer mass on the QC market equals

$$B_{\rm qc} = \widetilde{b}^* S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \frac{1}{\left(C_{\rm qa}\right)^{\frac{2}{1-\theta}}},\tag{A.6}$$

and the mass of non-participating buyers

$$\widetilde{b}_0 = \max\left\{B - \overline{b}\left(S + \widetilde{\mu}(S_{\rm qc}, C_{\rm qa})\right), 0\right\},\tag{A.7}$$

where $\bar{b} = \left(\frac{\lambda\rho\delta}{c}\right)^{\frac{2}{1-\theta}}$ is defined in Eq. (16),

$$\widetilde{b}^* = \min\left\{\frac{B}{S + \widetilde{\mu}(S_{\rm qc}, C_{\rm qa})}, \overline{b}\right\},\tag{A.8}$$

and

$$\widetilde{\mu}(S_{\rm qc}, C_{\rm qa}) = S_{\rm qc} \left(\left(\frac{S_{\rm qc}^{\theta}}{C_{\rm qa}} \right)^{\frac{2}{1-\theta}} - 1 \right). \tag{A.9}$$

Proof of Lemma A.2. If in equilibrium $S_{qc} = 0$, then the equilibrium reduces to the AS-only equilibrium described in Theorem 1. We can easily verify that all results hold.

Next, consider an equilibrium in which $S_{qc} > 0$. Because in equilibrium buyers across AS submarkets earn the same profit, we have that $\tilde{b}_j = \tilde{b}^*$ for any AS submarket *j*. Additionally, Lemma A.1 implies Eq. (A.6). Thus, the total mass of buyers

$$B \ge B - \widetilde{b}_0 = B_{\rm as} + B_{\rm qc} = \widetilde{b}^* S_{\rm as} + \widetilde{b}^* S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \frac{1}{(C_{\rm qa})^{\frac{2}{1-\theta}}} = \widetilde{b}^* \left(S - S_{\rm qc} + S_{\rm qc} \left(\frac{S_{\rm qc}^\theta}{C_{\rm qa}} \right)^{\frac{2}{1-\theta}} \right).$$
(A.10)

Thus,

$$\widetilde{b}^* \le \frac{B}{S - S_{\rm qc} + S_{\rm qc} \left(\frac{S_{\rm qc}^{\theta}}{C_{\rm qa}}\right)^{\frac{2}{1 - \theta}}} = \frac{B}{S + \widetilde{\mu}(S_{\rm qc}, C_{\rm qa})} \tag{A.11}$$

There are two possibilities depending on the mass of non-participating buyers \tilde{b}_0 .

- If $\tilde{b}_0 > 0$, then some buyers do not participate and earn zero profit, which implies that every participating buyer also earns zero profit ($\tilde{\psi}^b = 0$). Hence, the mass on each AS submarket $\tilde{b}^* = \bar{b}$ and $C_{qa} = c_{qc}/c$. Thus, based on Lemma A.1, $B_{qc} = \bar{b}S_{qc}^{\frac{1+\theta}{1-\theta}} \left(\frac{c}{c_{qc}}\right)^{\frac{2}{1-\theta}}$. It follows that $\tilde{b}_0 = B B_{as} B_{qc} = B \bar{b}\left(S + \tilde{\mu}(S_{qc}, C_{qa})\right)$. Because $\tilde{b}_0 > 0$, we have that $\tilde{b}^* = \bar{b} < \frac{B}{S + \tilde{\mu}(S_{qc}, C_{qa})}$.
- If $\tilde{b}_0 = 0$, all buyers participate. Hence, $B = B_{as} + B_{qc} = \tilde{b}^* (S + \tilde{\mu}(S_{qc}, C_{qa}))$, which implies that $\tilde{b}^* = \frac{B}{S + \tilde{\mu}(S_{qc}, C_{qa})}$. Buyers on AS submarkets earn non-negative profit, which implies that $\frac{B}{S + \tilde{\mu}(S_{qc}, C_{qa})} \leq \bar{b}$ and $B \leq \bar{b} (S + \tilde{\mu}(S_{qc}, C_{qa}))$.

Thus, we have Eqs. (A.7) and (A.8).

We need to solve the equation $B = \tilde{b}^* (S + \tilde{\mu}(S_{qc}, C_{qa}))$ to find equilibrium \tilde{b}^* and S_{qc} because C_{qa} , based on Eq. (A.2), generally depends on \tilde{b}_j when $c_{qc} \neq c$.

Proposition A.1. If $c \le c_{qc} \le cS_{qc}^{\theta}$, then in the parallel-trading equilibrium,

$$b_0 \ge \widetilde{b}_0,$$
 (A.12)

$$\widetilde{b}^* \le b^* \le B_{\rm qc}/S_{\rm qc},\tag{A.13}$$

$$\widetilde{\pi}_{as}^{s} \le \pi^{s} \le \widetilde{\pi}_{qc}^{s}, \tag{A.14}$$

$$\pi^b \le \widetilde{\pi}^b_{\rm as} \le \widetilde{\pi}^b_{\rm qc},\tag{A.15}$$

$$\psi^b \le \widetilde{\psi}^b, \tag{A.16}$$

$$\psi^{s} \ge \widetilde{\psi}^{s}_{as},$$
(A.17)

$$V \le \tilde{V},\tag{A.18}$$

$$\Omega \le \widetilde{\Omega}.\tag{A.19}$$

Proof of Proposition A.1. We first show that $c_{qc} \le cS_{qc}^{\theta}$ implies that $C_{qa} \le S_{qc}^{\theta}$. If $c_{qc} \le c$, then $C_{qa} = \frac{c_{qc} + \psi^b}{c + \psi^b} \le 1 \le S_{qc} \le S_{qc}^{\theta}$; if $c < c_{qc} \le c_{as}S_{qc}^{\theta}$, then $C_{qa} = \frac{c_{qc} + \tilde{\psi}^b}{c + \tilde{\psi}^b} \le \frac{c_{qc}}{c} \le S_{qc}^{\theta}$.

Based on Eq. (A.9), $C_{qa} \leq S_{qc}^{\theta}$ implies that $\tilde{\mu}(S_{qc}, C_{qa}) \geq 0$. Then, Eq. (A.8) implies that $\tilde{b}^* \leq \min\{\frac{B}{S}, \bar{b}\} = b^*$ and Eq. (A.7) implies that $\tilde{b}_0 \leq \max\{B - \bar{b}S, 0\} = b_0$. Because $B_{qc} = B - \tilde{b}_0 - B_{as} \geq B - b_0 - b^*(S - S_{qc})$, we have that $B_{qc} - b^*S_{qc} \geq B - b_0 - b^*S \geq B - b_0 - \bar{b}S \geq 0$, which implies that $B_{qc}/S_{qc} \geq b^*$.

Given that $\tilde{b}^* \leq b^* \leq B_{\rm qc}/S_{\rm qc}$, we have that $\tilde{\pi}^s_{\rm as} \leq \pi^s \leq \tilde{\pi}^s_{\rm qc}$ and $\pi^b \leq \tilde{\pi}^b_{\rm as}$. A buyer's profit $\tilde{\psi}^b = \tilde{\pi}^b_{\rm as}\rho\delta - c \geq \pi^b\rho\delta - c = \psi^b$. Hence, the total profit of all buyers $(B - \tilde{b}_0)\tilde{\psi}^b \geq (B - b_0)\psi^b$. The profit of an AS seller $\tilde{\psi}^s_{\rm as} = \tilde{\pi}^s_{\rm as}(1 - \rho)\delta \leq \pi^s(1 - \rho)\delta = \psi^s$.

Moreover, because every buyer earns the same profit $\tilde{\psi}^b$, $c_{qc} \ge c$ implies that $\tilde{\pi}^b_{qc} = \frac{c_{qc} + \tilde{\psi}^b}{\rho \delta} \ge \frac{c + \tilde{\psi}^b}{\rho \delta} = \tilde{\pi}^b_{as}$. Hence, $\tilde{\pi}^b_{qc} \ge \tilde{\pi}^b_{as} \ge \pi^b$. It follows that the total trading volume $\tilde{V} = B_{as}\tilde{\pi}^b_{as} + B_{qc}\tilde{\pi}^b_{qc} \ge (B_{as} + B_{qc})\pi^b \ge (B - b_0)\pi^b = V$. Thus, the average selling probability per asset $\tilde{\pi}^{avg} = \tilde{V}/S \ge V/S = \pi^{avg}$. The total profit of all sellers are higher because $S\tilde{\pi}^{avg}(1-\rho)\delta \ge S\pi^{avg}(1-\rho\delta)$. Thus, the welfare gains of all traders $\tilde{\Omega} \ge \Omega$.

B.2 Heterogeneous Trading Gain across Assets

In the main model, the trading of any asset yields the same trading gain of δ . In this section, we relax this assumption. Specifically, we assume asset *j* is worth v_j to sellers and $v_j + \delta_j$ to buyers, so the trading of asset *j* leads to a gain of δ_j . Consistent with empirical evidence, we also assume that δ_j weakly increases with asset value v_j .²⁸

In this section, we first derive the equilibria for the AS-only market (Proposition A.2) and the parallel-trading market (Proposition A.3) under asset heterogeneity. We then compare these equilibria and show in Proposition A.4 that, as in the main model, the parallel-trading market sees more buyer participation and worse liquidity for assets traded via AS trading. Moreover, assets traded via QC in the parallel-trading equilibrium are, *on average*, more liquid than in the AS-only equilibrium.

The equilibrium for the AS-only market is as follows.²⁹

Proposition A.2. In the AS-only equilibrium, the mass of buyers on AS submarket j equals

$$b_{j} = \begin{cases} b_{j}^{*} & \text{if } j \in \{1, 2, \cdots, S\}, \\ B - \sum_{k \in \mathscr{A}} b_{k}^{*} & \text{if } j = 0, \end{cases}$$
(A.20)

28

²⁹The equilibrium described in Proposition A.2 reduces to the one described in Theorem 1 when $\delta_j = \delta$ for every asset j because in this situation $\sum_{k \in \mathcal{A}} \delta_k^{\frac{2}{1-\theta}} = S \delta_j^{\frac{2}{1-\theta}} = S \delta_j^{\frac{2}{1-\theta}}$ and $\left(\frac{\lambda \rho \delta_j}{c}\right)^{\frac{2}{1-\theta}} = \left(\frac{\lambda \rho \delta}{c}\right)^{\frac{2}{1-\theta}}$.

where

$$b_{j}^{*} = \min\left\{B \cdot \frac{\delta_{j}^{\frac{2}{1-\theta}}}{\sum_{k \in \mathscr{A}} \delta_{k}^{\frac{2}{1-\theta}}}, \left(\frac{\lambda \rho \delta_{j}}{c}\right)^{\frac{2}{1-\theta}}\right\}$$
(A.21)

Proof of Proposition A.2. The buying probability on an AS submarket $\pi_j^b = \lambda/b_j^{\frac{1-\theta}{2}}$. In equilibrium, a buyer on any AS submarket earns the same profit. Hence, for any asset $j \in \mathcal{A}$, a buyer's profit $\psi^b = \pi_j^b \rho \delta_j - c = \frac{\lambda \rho \delta_j}{b_j^{\frac{1-\theta}{2}}} - c$, which implies that $b_j = \left(\frac{\lambda \rho}{c+\psi^b}\right)^{\frac{2}{1-\theta}} \delta_j^{\frac{2}{1-\theta}}$. It follows that

$$B \ge \sum_{k \in \mathscr{A}} b_k = \left(\frac{\lambda \rho}{c + \psi^b}\right)^{\frac{2}{1-\theta}} \sum_{k \in \mathscr{A}} \delta_k^{\frac{2}{1-\theta}}.$$
 (A.22)

Hence, $b_j = \left(\frac{\lambda\rho}{c+\psi^b}\right)^{\frac{2}{1-\theta}} \delta_j^{\frac{2}{1-\theta}} \le \frac{B}{\sum_{k \in \mathcal{A}} \delta_k^{\frac{2}{1-\theta}}} \delta_j^{\frac{2}{1-\theta}}$. Because $\psi^b \ge 0$, we have that $b_j \le \delta_j^{\frac{2}{1-\theta}} \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}}$. Therefore, $b_j \le b_j^*$. If $b_j < b_j^*$, then some buyers do not participate $(B > \sum_{k \in \mathcal{A}} b_k)$ and yet every participating buyer earns strictly positive profit $(\psi^b > 0)$, which cannot be an equilibrium. Therefore, $b_j = b_j^*$.

Proposition A.2 implies that assets with higher trading gains are more liquid $(\partial \pi_j^s / \partial \delta_j \ge 0)$. Intuitively, a higher trading gain attracts more buyers, making the asset easier to sell and more liquid. Additionally, because the trading gain δ_j increases with asset value v_j , more valuable assets experience better liquidity $(\partial \pi_i^s / \partial v_j \ge 0)$.

We then derive the parallel-trading equilibrium. In the QC market, the price P_{qc} equals $v_{qc} + \delta_{qc}$ with probability $1 - \rho$ and equals v_{qc} with probability ρ , where $\delta_{qc}^{\frac{2}{1-\theta}} =$

 $\mathbf{E}[\delta_{\tau}^{\frac{2}{1-\theta}} | \tau \in \widetilde{\mathcal{M}}_{qc}].^{30}$ Additionally, for ease of exposition, we define

$$\delta(v^*) = \delta_k \text{ when } v_k = v^*, \tag{A.23}$$

$$\bar{\delta}(v^*) := \left(\mathbf{E} \left[\left. \delta_k^{\frac{2}{1-\theta}} \right| k \in \mathscr{A} \text{ and } v_k \in [\underline{v}, v^*] \right] \right)^{\frac{1-\theta}{2}}, \tag{A.24}$$

$$\widetilde{\eta}(v^*, \delta_j) := \mathbf{E}\left[v|v \in [\underline{v}, v^*]\right] + (1-\rho)\widetilde{\delta}(v^*) \left(1 - \frac{\delta_j^{\frac{1}{2-\theta}}}{\widetilde{\delta}(v^*)^{\frac{2}{1-\theta}} \left(S \cdot \Pr\left\{v \in [\underline{v}, v^*]\right\}\right)^{\frac{2\theta}{1-\theta}}}\right). \quad (A.25)$$

The parallel-trading equilibrium is as follows.

Proposition A.3. In the parallel-trading equilibrium, sellers' venue choices and buyers' masses are as follows.

• The seller of asset j chooses QC trading if the asset's value $v_j \in [\underline{v}, \overline{v}]$ and AS trading if otherwise, where \overline{v} solves $\overline{v} = \min \{\eta(\overline{v}, \delta(\overline{v})), v_{\max}\}$. That is,

$$\widetilde{\mathcal{M}}_{qc} = \left\{ j : j \in \mathcal{A} \text{ and } \underline{v} \le v_j \le \overline{v} \right\},$$
(A.26)

$$\widetilde{\mathcal{M}}_{as} = \mathscr{A} - \widetilde{\mathcal{M}}_{qc}. \tag{A.27}$$

• The mass of buyers who participate in the QC market equals

$$B_{\rm qc} = \frac{S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}} \cdot \min\left\{B, \bar{B}\right\},\tag{A.28}$$

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

$$\widetilde{b}_{j}^{*} = \frac{\delta_{j}^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}}} \cdot \min\left\{B, \overline{B}\right\} \quad \text{for any } j \in \widetilde{\mathcal{M}}_{as}, \tag{A.29}$$

³⁰If δ_{qc} were the arithmetic mean of the trading gains of assets on the QC market, buyers would have less incentive to choose QC trading due to the Jensen's inequality effect because, as Eq. (A.21) shows, the mass of buyers b_j in an AS submarket is proportional to $\delta_j^{\frac{2}{1-\theta}}$. We assume that $\delta_{qc}^{\frac{2}{1-\theta}} = \mathbf{E}[\delta_{\tau}^{\frac{2}{1-\theta}} | \tau \in \widetilde{\mathcal{M}}_{qc}]$ to abstract away from this effect.

the total mass of buyers participating in all AS submarkets equals

$$B_{\rm as} = \frac{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_j^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_k^{\frac{2}{1-\theta}} + S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}} \cdot \min\{B, \bar{B}\}, \qquad (A.30)$$

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

$$\tilde{b}_0 = \max\left\{B - \bar{B}, 0\right\},\tag{A.31}$$

where

$$S_{\rm qc} = S \cdot \Pr\left\{v_j \in [\underline{v}, \bar{v}]\right\},\tag{A.32}$$

$$\bar{B} = \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} \left(\sum_{k\in\widetilde{\mathcal{M}}_{as}} \delta_j^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}}\right).$$
(A.33)

Proof of Proposition A.3. First, we derive buyer masses. In equilibrium, every buyer earns the same profit. Let $B_{as} = \sum_{k \in \widetilde{\mathcal{M}}_{as}} \widetilde{b}_k$ denote the total mass of buyers participating in AS trading. Then, Proposition A.2 implies that the mass of buyers on the AS submarket for asset j equals

$$\widetilde{b}_{j} = \delta_{j}^{\frac{2}{1-\theta}} \cdot \min\left\{\frac{B_{\mathrm{as}}}{\sum_{k \in \widetilde{\mathcal{M}}_{\mathrm{as}}} \delta_{k}^{\frac{2}{1-\theta}}}, \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}}\right\}.$$
(A.34)

In addition, an QC buyer expects to earn $\rho \tilde{\pi}^{b}_{qc} \delta_{qc} - c$ that equals an AS buyer's expected profit $\rho \tilde{\pi}^{b}_{as,j} \delta_{j} - c$. It follows that

$$\frac{\delta_{\rm qc}}{\delta_j} = \frac{\widetilde{\pi}^b_{\rm as,j}}{\widetilde{\pi}^b_{\rm qc}} = \frac{B_{\rm qc}^{\frac{1-\theta}{2}}}{S_{\rm qc}^{\frac{1-\theta}{2}}(\widetilde{b}_j)^{\frac{1-\theta}{2}}},\tag{A.35}$$

which implies that

$$B_{\rm qc} = S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}} \left(\frac{\widetilde{b}_j}{\delta_j^{\frac{2}{1-\theta}}} \right) = S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}} \min\left\{ \frac{B_{\rm as}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_k^{\frac{2}{1-\theta}}}, \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} \right\}.$$
(A.36)

There are two scenarios in terms of buyer profits.

• If each buyer earns zero profit, then

$$\widetilde{b}_{j} = \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} \delta_{j}^{\frac{2}{1-\theta}} = \frac{\delta_{j}^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}}} \cdot \overline{B},$$
(A.37)

$$B_{\rm qc} = \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}} = \frac{S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}} \cdot \bar{B}, \tag{A.38}$$

which implies that the total mass of participating buyers $\sum_{k \in \widetilde{\mathcal{M}}_{as}} \widetilde{b}_j + B_{qc} = \overline{B} \leq B$ and $\widetilde{b}_0 = B - \overline{B}$.

• If each buyer earns positive profit, then

$$\widetilde{b}_{j} = \delta_{j}^{\frac{2}{1-\theta}} \frac{B_{\rm as}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_{k}^{\frac{2}{1-\theta}}} < \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} \delta_{j}^{\frac{2}{1-\theta}}, \tag{A.39}$$

$$B_{\rm qc} = S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}} \frac{B_{\rm as}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_k^{\frac{2}{1-\theta}}} < \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}, \tag{A.40}$$

which implies that

$$B = \sum_{k \in \widetilde{\mathcal{M}}_{as}} \widetilde{b}_j + B_{qc} = B_{as} \left(1 + S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}} \frac{1}{\sum_{k \in \widetilde{\mathcal{M}}_{as}} \delta_k^{\frac{2}{1-\theta}}} \right) < \overline{B}.$$
(A.41)

Thus,

$$B_{\rm as} = \frac{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_k^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_k^{\frac{2}{1-\theta}} + S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}} \cdot B$$
(A.42)

$$\widetilde{b}_{j} = \frac{\delta_{j}^{\overline{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}}} \cdot B$$
(A.43)

$$B_{\rm qc} = \frac{S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}}{\sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{\rm qc}^{\frac{1+\theta}{1-\theta}} \delta_{\rm qc}^{\frac{2}{1-\theta}}} \cdot B, \tag{A.44}$$

$$\widetilde{b}_0 = B - B_{\rm as} - B_{\rm qc} = 0. \tag{A.45}$$

Therefore, Eqs. (A.28) to (A.31) and (A.33) hold in equilibrium.

Second, we derive a seller's venue choice taking other sellers' choices and buyer masses as given. Consider a seller of the asset with value $v_j \ge \underline{v}$. If $v_j \in [\underline{v}, \overline{v}]$, this seller chooses QC trading because in equilibrium no buyers participate in the AS submarket of this asset. If $v_j > \overline{v}$, then $\overline{v} < v_j \le v_{\text{max}}$, which, given the equilibrium condition that $\overline{v} = \min \{\eta(\overline{v}, \delta(\overline{v})), v_{\text{max}}\}$, implies that $v_{\text{max}} \ge v_j > \overline{v} = \widetilde{\eta}(\overline{v}, \delta(\overline{v}))$. The profit of selling asset *j* through AS trading equals $\widetilde{\psi}(j, v_j) = \widetilde{\pi}_{\text{as},j}^s (1 - \rho)\delta_j$ and that of selling it on the QC market equals $\widetilde{\psi}(\text{qc}, v_j) = \widetilde{\pi}_{\text{qc}}^s (v_{\text{qc}} - v_j + (1 - \rho)\delta_{\text{qc}})$. It follows that

$$\widetilde{\psi}(j,\nu_j) - \widetilde{\psi}(qc,\nu_j) = \widetilde{\pi}_{as,j}^s (1-\rho)\delta_j - \widetilde{\pi}_{qc}^s (\nu_{qc} - \nu_j + (1-\rho)\delta_{qc})$$
(A.46)

$$= \tilde{\pi}_{\rm qc}^{s} \left(\nu_j - \left(\nu_{\rm qc} + (1-\rho)\delta_{\rm qc} \left(1 - \frac{\delta_j \tilde{\pi}_{\rm as,j}^s}{\delta_{\rm qc} \tilde{\pi}_{\rm qc}^s} \right) \right) \right). \tag{A.47}$$

The buyer masses given in Eqs. (A.28) to (A.31) and (A.33) imply that

$$\frac{\delta_j}{\delta_{\rm qc}} \frac{\widetilde{\pi}_{\rm as,j}^s}{\widetilde{\pi}_{\rm qc}^s} = \frac{S_{\rm qc}^{\frac{1+\theta}{2}}(\widetilde{b}_j)^{\frac{1-\theta}{2}}}{B_{\rm qc}^{\frac{1-\theta}{2}}} \frac{S_{\rm qc}^{\frac{1-\theta}{2}}\widetilde{b}_j^{\frac{1+\theta}{2}}}{B_{\rm qc}^{\frac{1+\theta}{2}}} = \frac{S_{\rm qc}}{B_{\rm qc}} \widetilde{b}_j = \left(\frac{\delta_j}{\delta_{\rm qc}}\right)^{\frac{2}{1-\theta}} \frac{1}{S_{\rm qc}^{\frac{2\theta}{1-\theta}}}.$$
(A.48)

Thus,

$$\frac{\widetilde{\psi}(j,\nu_j) - \widetilde{\psi}(qc,\nu_j)}{\widetilde{\pi}_{qc}^s} = \nu_j - \left(\nu_{qc} + (1-\rho)\delta_{qc} \left(1 - \frac{\left(\frac{\delta_j}{\delta_{qc}}\right)^2}{\frac{2\theta}{1-\theta}}\right)\right) = \nu_j - \widetilde{\eta}(\bar{\nu},\delta_j).$$
(A.49)

Because $\delta(v_k)$ weakly increases with v_k and $v_j > \bar{v}$, we have that $\delta_j = \delta(v_j) \ge \delta(\bar{v})$. Hence, $\tilde{\eta}(\bar{v}, \delta_j) \le \tilde{\eta}(\bar{v}, \delta(\bar{v})) = \bar{v} < v_j$. Therefore, $v_j - \tilde{\eta}(\bar{v}, \delta_j) > 0$ and $\tilde{\psi}(j, v_j) > \tilde{\psi}(qc, v_j)$, so the seller should choose AS trading.

Next, we compare the AS-only equilibrium and the parallel-trading equilibrium and find the following results.

Proposition A.4. Comparing with the AS-only equilibrium, in the parallel-trading equilibrium, (1) more buyers participate $(B - \tilde{b}_0 \ge B - b_0)$, (2) each buyer earns more profit $(\tilde{\psi}^b \ge \psi^b)$, (3) assets traded on AS submarket experience worse liquidity ($\tilde{\pi}_{as,j}^s \le \pi_j^s \forall j \in \widetilde{\mathcal{M}}_{as}$), and (4) assets traded on the QC market on average experience better liquidity ($\tilde{\pi}_{qc}^s \ge \frac{1}{S_{qc}} \sum_{k \in \widetilde{\mathcal{M}}_{qc}} \pi_k^s$).

Proof of Proposition A.4. Because $S_{qc} \ge 1$, we have that $S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}} = S_{qc}^{\frac{2\theta}{1-\theta}} \sum_{k \in \widetilde{\mathcal{M}}_{qc}} \delta_k^{\frac{2}{1-\theta}} \ge \sum_{k \in \widetilde{\mathcal{M}}_{qc}} \delta_k^{\frac{2}{1-\theta}}$. Hence,

$$\sum_{k \in \widetilde{\mathcal{M}}_{as}} \delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}} \delta_{qc}^{\frac{2}{1-\theta}} \ge \sum_{k \in \widetilde{\mathcal{M}}_{as}} \delta_{k}^{\frac{2}{1-\theta}} + \sum_{k \in \widetilde{\mathcal{M}}_{qc}} \delta_{k}^{\frac{2}{1-\theta}} = \sum_{k \in \mathscr{A}} \delta_{k}^{\frac{2}{1-\theta}}.$$
 (A.50)

Then, Eqs. (A.33) and (A.50) imply that $\bar{B} \ge \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} \sum_{k \in \mathscr{A}} \delta_k^{\frac{2}{1-\theta}}$. It follows that

$$B - \tilde{b}_0 = \min\left\{B, \bar{B}\right\} \ge \min\left\{B, \left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}} \sum_{k \in \mathscr{A}} \delta_k^{\frac{2}{1-\theta}}\right\} = B - b_0.$$
(A.51)

Further, Eqs. (A.21), (A.29), (A.33) and (A.50) imply that for any asset j traded through

AS trading $(\forall j \in \widetilde{\mathcal{M}}_{as})$, we have that

$$\widetilde{b}_{j}^{*} = \min\left\{\frac{B\delta_{j}^{\frac{2}{1-\theta}}}{\sum_{k\in\widetilde{\mathcal{M}}_{as}}\delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}}\delta_{qc}^{\frac{2}{1-\theta}}}, \frac{B\delta_{j}^{\frac{2}{1-\theta}} \cdot}{\sum_{k\in\widetilde{\mathcal{M}}_{as}}\delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}}\delta_{qc}^{\frac{2}{1-\theta}}}\right\}$$

$$\leq \min\left\{\frac{B\delta_{j}^{\frac{2}{1-\theta}}}{\sum_{k\in\widetilde{\mathcal{A}}}\delta_{k}^{\frac{2}{1-\theta}}}, \frac{\left(\frac{\lambda\rho}{c}\right)^{\frac{2}{1-\theta}}\left(\sum_{k\in\widetilde{\mathcal{M}}_{as}}\delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}}\delta_{qc}^{\frac{2}{1-\theta}}\right)}{\sum_{k\in\widetilde{\mathcal{M}}_{as}}\delta_{k}^{\frac{2}{1-\theta}} + S_{qc}^{\frac{1+\theta}{1-\theta}}\delta_{qc}^{\frac{2}{1-\theta}}}\right\}$$

$$= \min\left\{B\frac{\delta_{j}^{\frac{2}{1-\theta}}}{\sum_{k\in\widetilde{\mathcal{A}}}\delta_{k}^{\frac{2}{1-\theta}}}, \left(\frac{\lambda\rho\delta_{j}}{c}\right)^{\frac{2}{1-\theta}}}{\sum_{k\in\widetilde{\mathcal{A}}}\delta_{k}^{\frac{2}{1-\theta}}}\right\} = b_{j}^{*}.$$
(A.52)

Hence,

$$\widetilde{\pi}_{\mathrm{as},j}^{b} = \lambda \left(\frac{1}{\widetilde{b}_{j}^{*}}\right)^{\frac{1-\theta}{2}} \ge \lambda \left(\frac{1}{b_{j}^{*}}\right)^{\frac{1-\theta}{2}} = \pi_{j}^{b}, \tag{A.53}$$

$$\widetilde{\pi}_{\mathrm{as},j}^{s} = \lambda \left(\widetilde{b}_{j}^{*}\right)^{\frac{1+\theta}{2}} \le \lambda \left(b_{j}^{*}\right)^{\frac{1+\theta}{2}} = \pi_{j}^{s}.$$
(A.54)

Hence, each buyer's profit

$$\widetilde{\psi}^{b} = \widetilde{\pi}^{b}_{\mathrm{as},j} \rho \delta_{j} - c \ge \pi^{b}_{j} \rho \delta_{j} - c = \psi^{b}.$$
(A.55)

Moreover, because $B - \tilde{b}_0 \ge B - b_0$ and $\tilde{b}_j^* \le b_j^*$ for any $j \in \widetilde{\mathcal{M}}_{as}$, we have that

$$B_{\rm qc} = B - \tilde{b}_0 - \sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} \tilde{b}_k^* \ge B - b_0 - \sum_{k \in \widetilde{\mathcal{M}}_{\rm as}} b_k^* = \sum_{k \in \widetilde{\mathcal{M}}_{\rm qc}} b_k^*.$$
(A.56)

It follows that $B_{qc}/S_{qc} \ge \sum_{k \in \widetilde{\mathcal{M}}_{qc}} b_k^*/S_{qc}$. In addition, $\theta < 1$. Hence,

$$\widetilde{\pi}_{qc}^{s} = \lambda S_{qc}^{\theta} \left(\frac{B_{qc}}{S_{qc}}\right)^{\frac{1+\theta}{2}} \ge \lambda S_{qc}^{\theta} \left(\frac{\sum_{k \in \widetilde{\mathcal{M}}_{qc}} b_{k}^{*}}{S_{qc}}\right)^{\frac{1+\theta}{2}} \ge \lambda S_{qc}^{\theta} \frac{\sum_{k \in \widetilde{\mathcal{M}}_{qc}} \left(b_{k}^{*}\right)^{\frac{1+\theta}{2}}}{S_{qc}}$$
(A.57)

$$\geq \lambda \frac{\sum_{k \in \widetilde{\mathcal{M}}_{qc}} (b_k^*)^{\frac{1}{2}}}{S_{qc}} = \frac{1}{S_{qc}} \sum_{k \in \widetilde{\mathcal{M}}_{qc}} \pi_k^s.$$
(A.58)

Although QC assets are more liquid *on average*, there could exist an QC asset that is less liquid in the parallel-trading equilibrium.

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," *The Journal of Finance*, 2022, 77 (6), 3249–3287.
- **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.

IA.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. 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 highbalance 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 Theorem 2, varying \underline{v} changes the proportion of assets traded on the QC market q (defined in Eq. (47)) both by directly determining the set of assets that are excluded from QC trading $[v_{\min}, \underline{v})$ and by indirectly affecting the fraction of QC-eligible assets that would be traded via QC contracts as captured by $[\underline{v}, \overline{v}]$. 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 \overline{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 \overline{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 \overline{v} , its effects on the proportion of assets traded on the QC market q and on market liquidity are ambiguous.³¹ 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)."³² 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 \overline{v}_1 to \overline{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)

³¹Through numerical analysis, we have verified that both scenarios may occur.

³²Available at https://www.sifma.org/resources/submissions/testimony-on-behalf-of-sifma-at-hcfshearing-on-conforming-loan-limit-increase-and-impact-on-homebuyers-and-housing-market/.



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

assets in group ③ remain on the QC market, and (4) assets in group ④ switch from AS submarkets to the QC market.³³ 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 <u>v</u> 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. (48), \tilde{b}^* decreases. It follows that $\tilde{\pi}^b$ increases and $\tilde{\pi}^s_{as}$ decreases. In addition, Eq. (46) implies that \tilde{b}_0 decreases, so the mass of participating buyers $B - \tilde{b}_0$ increases. Thus, total trading volume $\tilde{V} = \tilde{\pi}^b (B - \tilde{b}_0)$ increases and social welfare $\tilde{\Omega} = (B - \tilde{b}_0)(\tilde{\pi}^b \delta - c)$ increases. \Box

³³By contrast, as we show in Section 4, only two subsets of assets exist when QC trading is added to a market with only AS trading: assets that remain on AS submarkets and assets that switch from AS submarkets to the QC market.

That is, the effect on liquidity varies across assets. First, liquidity deteriorates for assets that remain on AS submarkets because, as Eq. (48) shows, the mass of buyers on each AS submarket \tilde{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. (57), 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_{qc}/S_{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_{qc}/S_{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 \tilde{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.³⁴

³⁴Note that raising \underline{v} could lower q, which would lead to worse overall liquidity. Thus, it is necessary to empirically verify the premise that limiting super-conforming loans raises q.



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

IA.2 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}_{qc} and accept the ask price $\hat{P}_{qc} + \delta$, where

$$\hat{P}_{qc} = \zeta \cdot \underline{v} + (1 - \zeta) \mathbf{E}[v_k | k \in \widetilde{\mathcal{M}}_{qc}].$$
(IA.59)

When buyers' risk attitude $\zeta = 0$, buyers are risk-neutral; when $\zeta = 1$, buyers are ambiguityaverse 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 ($\overline{v} \downarrow$). 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 $\widetilde{\pi}_{as}^{s}$. Because fewer traders use QC trading, the total trading volume \widetilde{V} declines.