STAFF REPORTS

NO. 931
JULY 2020

REVISED
OCTOBER 2021

Asset Pricing with Cohort-Based Trading in MBS Markets

Nicola Fusari | Wei Li | Haoyang Liu | Zhaogang Song

FEDERAL RESERVE BANK of NEW YORK

Asset Pricing with Cohort-Based Trading in MBS Markets
Nicola Fusari, Wei Li, Haoyang Liu, and Zhaogang Song
Federal Reserve Bank of New York Staff Reports, no. 931
July 2020; revised October 2021
JEL classification: G12, G18, G21, E58

Abstract

Agency MBSs with diverse characteristics are traded in parallel through individualized specified pool (SP) contracts and standardized to-be-announced (TBA) contracts with delivery flexibility. This parallel trading environment generates distinctive effects on MBS pricing and trading: (1) Although cheapest-to-deliver (CTD) issues are present in TBA trading and absent from SP trading by design, MBS heterogeneity associated with CTD discounts affects SP yields positively, with the effect stronger for lower-value SPs; (2) high selling pressure amplifies the effects of MBS heterogeneity on SP yields; (3) greater MBS heterogeneity dampens SP and TBA trading activities but increases their ratio.

Key words: cohort, heterogeneity, liquidity, MBS, prepayment, TBA

Liu: Federal Reserve Bank of New York (email: haoyang.liu@ny.frb.org). Fusari, Li, Song: The Johns Hopkins Carey Business School (emails: nfusari1@jhu.edu, wei.li@jhu.edu, zsong8@jhu.edu). The authors thank Doug McManus at Freddie Mac and Marty Young at Goldman Sachs for helpful conversations on the MBS market. For helpful comments, they thank Wei Xiong (the editor), the associate editor, two anonymous referees, Kenneth Ahern, Nina Boyarchenko, Chris Jones, Francis Longstaff, John Miller, and seminar participants at Freddie Mac, Johns Hopkins Carey, UIC, and USC Marshall. David Rubio provided outstanding research assistance.

The market for agency mortgage-backed securities (MBSs), guaranteed by Fannie Mae, Freddie Mac, and Ginnie Mae, is one of the largest fixed-income markets in the U.S., with an outstanding amount of about \$8.8 trillion as of December 2019 according to the Securities Industry and Financial Markets Association (SIFMA). Agency MBSs are among the most important liquid and safe assets, receiving a low haircut in the liquidity coverage ratio requirement of Basel III and accounting for a large fraction of the tri-party repo collateral (He and Song (2021)). The agency MBS market has also played a prominent role in the implementation of the U.S. monetary policy since the global financial crisis through multiple rounds of quantitative easing, and the Federal Open Market Committee plans to keep involving agency MBSs in its regular policy operations (Frost et al. (2015); FRBNY (2020)).[1]

The remarkable liquidity of agency MBSs is often attributed to the market structure of trading. In particular, agency MBSs are traded via two parallel mechanisms: (1) specified pool (SP) trading, in which individual MBSs are traded using specific contracts and (2) to-be-announced (TBA) trading, in which similar (but nonidentical) MBSs are traded at the same price using a standardized contract. A TBA contract specifies, for example, only that a delivered MBS must be guaranteed by

[1]In response to the COVID-19 crisis, for example, the Federal Reserve announced purchases of agency MBSs together with Treasury securities (see Chen et al. (2021) for details).

Fannie Mae, consist of 30-year fixed-rate mortgages, and pay a coupon of 4% interest, usually known as a *coupon cohort*. By combining thousands of heterogeneous MBSs into a consolidated cohort, TBA trading incurs low transaction cost and serves as the bedrock of market liquidity of the whole agency MBS market (Gao et al. (2017); Li and Song (2020)). Such cohort trading with delivery flexibility is also used in other markets, such as Treasury futures, commodity futures, and general collateral repo contracts, and has been advocated to improve liquidity in other fixed-income markets, such as corporate bonds and municipal bonds.[2]

Given the close relationship between liquidity and asset pricing, it is somewhat surprising that the effects of the parallel trading environment on MBS pricing has received little attention. Existing studies on MBS pricing mostly focus on prepayment risks resulting from the uncertain timing of cash flows.[3] In this paper, we show that the parallel

[2]Bessembinder et al. (2019), for example, ask whether there is "scope for the trading of packages of corporate bonds based on a set of prescribed characteristics." 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."

[3]The timing of cash flows is uncertain because mortgage borrowers can prepay without penalty, and would do so particularly when interest rates decline. See, for example, the recent contributions by Gabaix et al. (2007), Chernov et al. (2017), Boyarchenko et al. (2019), and Diep et al. (2021).

trading environment influences MBS pricing and trading through distinctive economic channels, resulting in large pricing variations on top of those driven by prepayment risks. Our findings imply that the sound functioning of parallel trading is vital for the status of agency MBSs as liquid and safe assets. The economic channels we document also shed light on the potential effects of introducing a cohort trading mechanism to other markets.

To guide our empirical analyses, we propose a simple model to demonstrate the economic channels through which the parallel trading environment affects the trading and pricing of agency MBSs. In our model, heterogeneous MBSs with varying fundamental values are traded in two rounds before maturity. In trading round 1, all MBSs are sold; in trading round 2, some MBS owners experience liquidity shocks, forcing them to sell their MBSs. Sellers face a trade-off when choosing between TBA and SP trading in both trading rounds. On the one hand, transaction costs are higher in the SP market than in the TBA market, which is consistent with empirical evidence (the difference is 20-60 basis points according to Bessembinder et al. (2013) and Gao et al. (2017)). On the other hand, in the TBA market, because a single price is set for any MBS satisfying eligibility requirements, sellers have incentives to deliver the cheapest eligible MBSs. Given sellers' cheapest-to-deliver (CTD) option, buyers in the TBA market rationally bid prices that are lower than the average fundamental values of all eligible MBSs, resulting in CTD price discounts to MBSs traded on the TBA market. Such discounts are absent in the SP market because every MBS is priced individually.

The parallel trading environment influences MBS markets via two distinctive economic channels. First, higher-value MBSs are more likely to be sold on the SP market. Intuitively, if sellers of these MBSs use the TBA market, they would have to accept deeper CTD discounts because a single TBA price is set for any delivered MBS. We call this static effect the *venue selection* channel.[4] Second, when buyers bid for SP MBSs today, they take into account the potential costs of reselling these MBSs in the future. Because these buyers can use the TBA market as a backup selling venue when the SP market is illiquid in the future, SP prices today depend on the expected future CTD discounts in the TBA market.[5] We call this dynamic effect the *venue backup* channel.

MBS heterogeneity—the difference in value between the cheapest and the average MBS within a coupon cohort—affects MBS trading and pricing via both channels. First, when MBS heterogeneity is greater

[4]Downing et al. (2009) show that MBSs backing up collateralized mortgage obligation deals are less valuable than others, similar to the venue selection between TBA and SP markets.

[5]As put in Gao et al. (2017), the existence of the TBA market gives "potential buyers of an SP an option to deliver the SP in a TBA trade if market conditions change" in the future.

at a particular moment, CTD discounts in the TBA market are deeper, prompting more sellers to choose the SP market at the moment. Second, when expected *future* MBS heterogeneity is greater, buyers lower their bid prices in the SP market *today* because they expect deeper future CTD discounts, which makes the future TBA market a less valuable backup selling venue for them.

To empirically measure MBS heterogeneity, we use the difference in prepayment characteristics between the cheapest and average MBSs within a coupon cohort. Specifically, for the period from June 2003 through December 2018, we obtain monthly series of weighted-average original FICO scores (WAOCS), a key input to most MBS prepayment models, for all outstanding Fannie Mae 30-year MBSs as of each month. Higher WAOCS are associated with higher prepayment *risks* and lower MBS values (Fabozzi and Mann (2011)).[6] For each coupon cohort in every month, we measure MBS heterogeneity as the difference between the 95th percentile and the median of WAOCS, denoted as h^{WAOCS}, among the set of TBA-eligible MBSs.[7] Regressing h^{WAOCS} on its lagged

[6]From investors' perspective, MBSs with higher WAOCS are less valuable because borrowers with higher credit scores prepay more optimally: they are more likely to refinance when interest rate falls and less likely to do so when interest rate increases.

[7]All our main results remain robust to using other relevant MBS characteristics (e.g. the weighted average original loan size (WAOSIZE)) or other percentiles (e.g. the 90th percentile). To avoid potential outliers, we do not use the 100th percentile. We also construct a heterogeneity measure that combines different characteristics and

values delivers positive and highly significant coefficients, showing that the heterogeneity in WAOCS in the current period captures the expected future heterogeneity in prepayment rates reasonably well.

Using h^{WAOCS} as a measure of MBS heterogeneity, we test three main empirical hypotheses regarding the impact of the parallel trading environment on the pricing and trading of agency MBSs.

First, although the CTD issue is absent from SP trading (and present in TBA trading) by design, MBS heterogeneity associated with the CTD discount affects yields of SP MBSs positively through the venue backup channel. In particular, when MBS heterogeneity is greater, the TBA market as a future backup selling venue is less valuable to today's SP buyers, who then demand higher yields as compensation. Further, because of the venue selection channel, the effect of MBS heterogeneity on SP yields is weaker for more valuable SP MBSs because they are less likely to be sold on the TBA market in the future. These effects of MBS heterogeneity on SP yields reflect the distinctive impact of the parallel trading environment on pricing. In contrast, the dependence of TBA prices on MBS heterogeneity simply reflects the CTD discount embedded in TBA contracts.[8]

a heterogeneity measure based on realized prepayment rates, both of which deliver results similar to the baseline results. See Section IA.4 of the Internet Appendix for details.

[8]Early studies have examined the CTD discounts in futures contracts, including

We hence focus on testing the effects of MBS heterogeneity on SP yields in our main analyses. We follow Gabaix et al. (2007), Boyarchenko et al. (2019), and Song and Zhu (2019) to measure the MBS yield with the option-adjusted spread (OAS). Importantly, within each coupon cohort, we consider multiple groups of SP MBSs with distinct loan-to-value (LTV) ratios.[9] Consistent with our hypotheses, for SP MBSs with loan-to-value (LTV) ratios in the 80%-90% range, which are likely to be delivered into TBA contracts, a one-standard-deviation increase in h^{WAOCS} across coupon cohorts is associated with an increase in the OAS of about 17 basis points. The effect decreases to about 10 basis points for SP MBSs with LTV ratios in the 100%-105% range, which are eligible but less likely to be delivered into TBA contracts. In contrast, the effect is insignificant for TBA-ineligible SP MBSs. We also show that h^{WAOCS} positively affects the OAS of TBA MBSs, consistent with the presence of CTD discount.

Second, the effects of MBS heterogeneity on yields of SP MBSs are amplified when future liquidity shocks are more likely to occur: today's buyers of SP MBSs are more likely to sell these MBSs on the TBA mar-

---

e.g. Hegde (1988), Hemler (1990), Kane and Marcus (1986), and Gay and Manaster (1984), among others.

[9]Using the MBSs with fixed characteristics avoids the potential confounding issue when using the average of all SP MBSs—that the change in the average price of all SPs may simply reflect the change in the composition of MBSs sold on the SP market.

ket in the future, making SP yields more sensitive to MBS heterogeneity. We use the *Distress* measure of He et al. (2019),[10] which captures the "constrained" investment capital of large financial intermediaries who are major MBS investors. Panel regressions of the OAS on the interaction term of h^{WAOCS} with *Distress* generate positive and highly significant coefficients, confirming the stronger effects of MBS heterogeneity on SP yields during periods of heavy selling pressure.

Third, we test the effects of MBS heterogeneity on trading activities. Intuitively, an increase in MBS heterogeneity raises the CTD price discount and the effective costs of TBA trading, which in turn raises the effective costs of SP trading because owners of SP MBSs use TBA market as a backup. In consequence, trading activities should decline on both the TBA and SP markets. Further, deeper CTD price discounts would make sellers more willing to use the SP market relative to the TBA market, thereby increasing the ratio of SP to TBA trading activities. We empirically confirm both effects using MBS transaction data from the Financial Industry Regulatory Authority (FINRA) through its Trade Reporting and Compliance Engine (TRACE) that became available in May 2011. In particular, we find that a one-standard-deviation increase in h^{WAOCS} across coupon cohorts is associated with a decrease

[10]The *Distress* measure of He et al. (2019) is the first principal component of the balance-sheet-based leverage ratio measure of the aggregate intermediary sector of He et al. (2017) and the market-price-based "noise" measure of Hu et al. (2013).

of about \$62 billion and \$4 billion in TBA and SP monthly trading volume, respectively, as well as an increase of about 138% in the ratio of SP volume to TBA volume.

Our main results remain significant after controlling for potential misspecifications of the prepayment models that produce OAS measures, using alternative measures of MBS heterogeneity, using alternative samples, and using OAS based on the Libor swap curve. Further, we perform two analyses that distinguish the effects of MBS heterogeneity from those of prepayment risks. The first analysis hinges on the findings of Gabaix et al. (2007) and Diep et al. (2021) that the market price of prepayment risk shows opposite signs depending on whether premium or discount securities dominate the MBS market. We find, however, that the impact of MBS heterogeneity is positive regardless of whether premium or discount securities dominate. The second analysis examines prepayment risks of individual MBSs. Boyarchenko et al. (2019), for example, estimate the component of the non-interest-rate prepayment risk premium in the OAS by exploiting the fact that interest-only (IO) and principal-only (PO) MBS strips have opposite exposures to prepayment risks. We find, however, that MBS heterogeneity positively affects yields of both IO and PO strips, confirming that our heterogeneity measure is not a proxy for prepayment risk.

One may wonder whether it is worth studying the economic effects associated with the TBA/SP parallel trading environment because the

TBA market accounts for the majority of the MBS trading volume and the SP market appears tiny (Gao et al. (2017)). Note, however, that a substantial fraction of the TBA trading volume arises from investors' hedging and speculation activities that are often reversed before settlement and do not result in actual deliveries of MBSs. In fact, a rough estimate in An et al. (2020) shows that slightly more than half of newly issued TBA-eligible MBSs are actually sold through SP trading. Hence, the SP market is no less important than the TBA market insofar as facilitating mortgage loan securitization and reducing mortgage borrowers' costs. Furthermore, for coupon cohorts involving seasoned MBSs, the SP trading volume is actually larger than the TBA trading volume (see Table III).[11]

Our paper contributes to the asset pricing literature on MBS markets, most studies in which focus on prepayment risks. Early studies proposed valuation frameworks based on the contingent claims approach and econometric prepayment models.[12] Recently, Levin and

[11] In addition, TBA-eligible MBSs, which we focus on, make up the bulk of outstanding MBSs. TBA-ineligible MBSs, which are usually backed by high-balance mortgages, forty-year mortgages, and interest-only mortgages, account for less than 1% according to estimates of An et al. (2020).

[12] The contingent claims valuation framework is used in Dunn and McConnell (1981), Brennan and Schwartz (1985), Stanton and Wallace (1998), Dunn and Spatt (2005), Stanton (1995), Boudoukh et al. (1997), Titman and Torous (1989), Downing et al. (2005), and Longstaff (2005) among other studies. Studies based on economet-

Davidson (2005) and Boyarchenko et al. (2019) study implied prepayments of individual MBSs, while Chernov et al. (2017) study market-level implied prepayment factors by imposing no-arbitrage restrictions across MBSs. Moreover, Gabaix et al. (2007) and Diep et al. (2021) study the prepayment risk premium under a limits-of-arbitrage framework, while Duarte et al. (2007) document characteristics of various MBS portfolio strategies.[13]

Our paper is related in particular to studies that document the existence of a liquidity premium for MBSs. For example, Krishnamurthy and Vissing-Jorgensen (2013) and He and Song (2021) present evidence on the existence of scarcity premium and convenience premium for agency MBSs, while Bartolini et al. (2011) and Song and Zhu (2019) examine the premium of MBS as collateral in funding markets. Differing from these papers, ours shows that the parallel trading environment affects MBS pricing through distinctive economic channels.

In this regard, our paper is also related to the literature on MBS market structure and liquidity, including Bessembinder et al. (2013), Friewald et al. (2017), Gao et al. (2018), Schultz and Song (2019), Kim

---

ric prepayment models include Schwartz and Torous (1989), Richard and Roll (1989), and Deng et al. (2000). The prepayment model framework has been extended by Chen (1996) and Cheyette (1996) to estimate implied prepayments from MBS prices.

[13]Relatedly, Duarte (2007), Malkhozov et al. (2016) and Hansen (2014) study the effects of mortgage-risk hedging on Treasury and broader interest rate markets.

and Huh (2019), and Liu et al. (2021). Our paper adds to this literature by connecting MBS market microstructure to asset pricing, along the lines of the seminal work of Amihud and Mendelson (1986) and the literature surveyed in Easley and O'Hara (2003), Amihud et al. (2006) and Vayanos and Wang (2013).

The rest of the paper is organized as follows. In Section I, we introduce institutional background of the MBS market. In Section II, we present the stylized model. We describe our data in Section III and present main empirical results in Section IV. We conclude in Section V. Additional results are provided in the Internet Appendix.

I. Institutional Background

We provide a brief introduction to the agency MBS market, highlighting its unique trading environment (see Vickery and Wright (2013) and Gao et al. (2017) for additional details). Most agency MBSs are issued as pass-through securities in which interest payments (subtracting credit guarantee and mortgage service fees) and principal payments on underlying mortgages are passed through pro rata to MBS investors. Pass-through securities can be pooled together to create structured MBSs, such as collateralized mortgage obligations (CMOs) and interest-only and principal-only Separate Trading of Registered Interest and Principal of Securities (STRIPs). The structured MBSs create customized prepayment and maturity profiles by carving up mortgage cash flows. According to SIFMA, the outstanding balances of pass-through and structured MBSs are about $7.3 and $1.1 trillion, respectively. We focus mainly on pass-through MBSs, but also use STRIPs to distinguish the liquidity premium from the prepayment risk premium.

All agency MBSs are effectively default-free, with credit guarantees provided by Fannie Mae, Freddie Mac, or Ginnie Mae. They are, however, subject to uncertainty on the timing of cash flows, known as prepayment risk, because mortgage borrowers can prepay mortgage loans whenever they want. For example, when mortgage rates declines, increased refinancing activities will lead to earlier principal payments; in consequence, MBS investors receive larger cash flows that they can only invest for lower rates. MBSs differ substantially in prepayment risk because each MBS is "unique in its prepayment characteristics" (Gao et al. (2017)). This heterogeneity originates from the vastly different characteristics of mortgage loans and their borrowers (see Section III for summary statistics of different prepayment speeds of varying MBSs).

One might conjecture, given the large asset heterogeneity and OTC nature of trading, that the agency MBS market would be very illiquid, just like the corporate and municipal bond markets (Bessembinder et al. (2019)). On the contrary, a large portion of agency MBSs are traded through TBA contracts at low transaction costs of about 2 basis points, comparable to the trading costs in the U.S. Treasury market.

A TBA contract specifies a set of eligible securities (e.g. Fannie Mae 30-year fixed-rate MBSs with a 4% security coupon rate) and fixes a single price, but the particular MBS a seller delivers needs to be specified only two days before the settlement day.[14] As mentioned in Gao et al. (2017) and Bessembinder et al. (2019) and theoretically modeled by Li and Song (2020), by combining thousands of heterogeneous MBSs into a consolidated cohort, TBA contracts promote network externality and create substantial market liquidity. Nonetheless, the single cohort-level price for heterogeneous MBSs leads naturally to a CTD issue and results in price discounts for TBA MBSs. Intuitively, the TBA price discount relates positively to the cross-sectional dispersion of MBS values within a cohort, and negatively affects the liquidity-creation value of the TBA mechanism.

Agency MBSs are also traded on the parallel SP market, where buyers and sellers agree to exchange a particular MBS. MBSs that are ineligible for delivery into TBA contracts, such as those with an LTV ratio above 1.05 or with more than 10% of its pool value in jumbo-conforming loans, can be traded only as SP MBSs (Vickery and Wright (2013)). In-

[14]SIFMA sets eligibility criteria for TBA delivery and specifies settlement days. Details on these regulations 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.

stead, TBA-eligible MBSs can be traded on both the TBA and SP markets. Naturally, those with the most desirable prepayment characteristics are traded on the SP market because sellers can realize the full value of their MBSs rather than the TBA price with a CTD discount. In consequence, SP prices are usually quoted at a “pay up” relative to TBA prices. SP trading, however, incurs transaction costs that are about 20-60 basis points higher. Sellers of TBA-eligible MBSs hence face a tradeoff between the CTD price discount in the TBA market and the high trading cost in the SP market.

In addition to creating outright liquidity, TBA trading also improves liquidity of the parallel SP trading. Indeed, as shown by Gao et al. (2017), transaction cost declines sharply at the threshold of TBA eligibility. TBA trading can benefit SP trading through at least two channels. First, TBA trading allows investors to hedge their SP holdings. Second, TBA trading also serves as a “backup” option for SP holders to offload their MBSs quickly, when market conditions change or they experience balance-sheet constraints. Overall, TBA trading serves as the foundation of market liquidity across the entire MBS market.

II. Model and Testing Hypotheses

In this section, we first develop a simple model that demonstrates the economic effects of the TBA/SP parallel trading environment on MBS pricing and trading. The novel effects result from a dynamic channel of parallel trading: when traders bid for MBSs on the SP market today, they take into account the potential costs of reselling these MBSs on the TBA market in the future. Hence SP prices today depend on the expected future TBA transaction costs, which originate from CTD price discounts. Guided by the model, we set up the hypotheses for empirical testing.

A. A Simple Model of MBS Trading and Pricing

We abstract prepayment risk away from the model and focus on how the parallel trading environment affects the trading and pricing of MBSs that are eligible for trading in both the TBA market and the SP market.

The specific model setup is as follows. We normalize the time discount rate at zero. MBSs are traded at time 1 and 2 and mature at time 3. At time 1, all MBSs are sold. At time 2, a fraction ρ of MBS owners experience idiosyncratic liquidity shocks, forcing them to sell their MBSs. When a trader buys an MBS at time 1, she knows that, with probability ρ, she might have to sell the MBS at time 2 rather than holding it to maturity at time 3. The time-3 payoff of an MBS falls in the range $[v_m - h_d, v_m + h_u]$, where v_m is the median MBS payoff and is assumed to be fixed. We measure (downside) MBS heterogeneity with h_d, the difference in value between the median and the cheapest MB-

Ss. The measure conveniently captures the cross-sectional dispersion of MBS values that is relevant for TBA trading.[15]

We assume no transaction costs in the TBA market as a normalization, reflecting the much lower trading cost of TBA trading than SP trading (Bessembinder et al. (2013); Gao et al. (2017)). Because TBA contracts do not fix specific MBSs to be delivered, buyers expect sellers to deliver the cheapest eligible MBSs they have for a price P_t^{TBA} at time $t \in \{1,2\}$. This is the CTD issue in the TBA market, which embodies the "lemon's problem" described by Akerlof (1970). We assume, for simplicity, that TBA buyers recognize the CTD issue and bid

$$P_t^{\text{TBA}} = v_m - h_d. \tag{1}$$

This simplifying assumption enables us to capture in a tractable manner the impact of MBS heterogeneity h_d on TBA prices resulting from the CTD issue. When a trader sells an MBS with value v_k on the TBA market, she suffers a price discount of $v_k - P_t^{\text{TBA}}$, which equals $v_k - v_m + h_d$ and increases with MBS heterogeneity h_d (relative to the fixed v_m).

If a seller chooses the SP market, she must specify the identity of the MBS she intends to deliver. Every seller in the SP market needs

[15]The upside MBS heterogeneity measure h_u is irrelevant for TBA trading because MBSs of highest values are sold in the SP market. We also assume, for simplicity, that h_d stays constant over time, so h_d also represents the expected future heterogeneity.

to pay a cost C_t^{SP} to locate a buyer. Empirical studies, including Gao et al. (2017), find that SP transaction costs may fluctuate considerably depending on market conditions. We assume that before buyers bid and sellers choose the selling venue at time 1, they observe the current transaction cost C_1^{SP} and believe that C_2^{SP}, the future transaction cost at time 2, follows a simple two-point distribution,

$$C_2^{\mathrm{SP}} = \begin{cases} c_{2,h} & \text{with probability } \pi_h, \\ c_{2,\ell} & \text{with probability } 1-\pi_h, \end{cases} \tag{2}$$

where $c_{2,h} \geq c_{2,\ell} \geq 0$. At time 2, sellers choose the selling venue after observing C_2^{SP}.

We find the equilibrium using backward induction. We assume for simplicity that buyers in the SP market earn zero profits in expectation. Hence, at time 2, SP buyers bid

$$P_2^{\mathrm{SP}}(v_k) = v_k \tag{3}$$

for MBS with value v_k because every MBS will pay its fundamental value at time 3.[16] Sellers' time-2 strategy is as follows.

[16]We assume that traders agree on the value of any particular MBS for simplicity. In practice, because Fannie Mae and Freddie Mac publicly provide key characteristics of every agency MBS to all traders, information asymmetry between MBS traders is unlikely to be severe. Moderate level of information asymmetry may still arise for two

PROPOSITION 1 (Time-2 equilibrium): *Consider a trader who sells an MBS with value v_k at time 2. The trader sells the MBS on the TBA market at price $P_2^{\mathrm{TBA}} = v_m - h_d$ if $v_k \leq \bar{v}_2$ and on the SP market at price $P_2^{\mathrm{SP}}(v_k) = v_k$ if $v_k > \bar{v}_2$, where the threshold*

$$\bar{v}_2 := P_2^{\mathrm{TBA}} + C_2^{\mathrm{SP}} = v_m - h_d + C_2^{\mathrm{SP}}. \tag{4}$$

A seller chooses the less costly selling venue. If $v_k > \bar{v}_2$, she chooses the SP market because the CTD price discount in the TBA market for this MBS $v_k - P_2^{\mathrm{TBA}}$ exceeds the SP selling cost C_2^{SP}. Otherwise she chooses the TBA market. Because C_2^{SP} is random, the time-2 TBA value threshold equals

$$\bar{v}_2 = \begin{cases} \bar{v}_{2,h} & \text{with probability } \pi_h, \\ \bar{v}_{2,\ell} & \text{with probability } 1 - \pi_h, \end{cases} \tag{5}$$

---

reasons: First, MBS issuers who securitize loans into MBSs possess additional loan-level information not disclosed to Fannie and Freddie. Second, traders may differ in expertise in valuating MBSs. The SP trading cost C_t^{SP} in our model could reflect, in a reduced-form manner, the impact of such information asymmetry on SP trading. Because TBA contracts are standardized and TBA trading is more transparent, we expect that such information asymmetry affects TBA trading to a lesser degree.

where

$$\bar{v}_{2,h} := v_m - h_d + c_{2,h} \quad \text{and} \quad \bar{v}_{2,\ell} := v_m - h_d + c_{2,\ell}. \tag{6}$$

Ascertaining the SP price at time 1 is less straightforward. Because a trader who buys an MBS on the SP market at time 1 might be forced to sell it at time 2, the trader bids a price that is equal to the MBS's terminal payoff less its expected effective selling cost at time 2, which depends on the MBS's value because the MBS may be sold on the TBA market or the SP market.

Specifically, because a low-value MBS ($v_k < \bar{v}_{2,\ell}$) will always be sold through the TBA market and a high-value MBS ($v_k > \bar{v}_{2,h}$) will always be sold through the SP market at time 2, the effective selling cost equals $v_k - P_2^{\mathrm{TBA}} = v_k - v_m + h_d$ for a low-value MBS and C_2^{SP} for a high-value one. In contrast, an medium-value MBS ($\bar{v}_{2,\ell} \le v_k \le \bar{v}_{2,h}$) will be sold through the TBA market if the high SP cost $c_{2,h}$ is realized and through the SP market if the low SP cost $c_{2,\ell}$ is realized at time 2. In consequence, the expected effective selling cost of medium-value MBSs is the probability-weighted average of the TBA cost $v_k - v_m + h_d$ and the SP cost $c_{2,\ell}$. These results are formalized as follows.

LEMMA 1 (Time-1 SP price): *At time* 1*, buyers in the SP market are* *willing to pay*

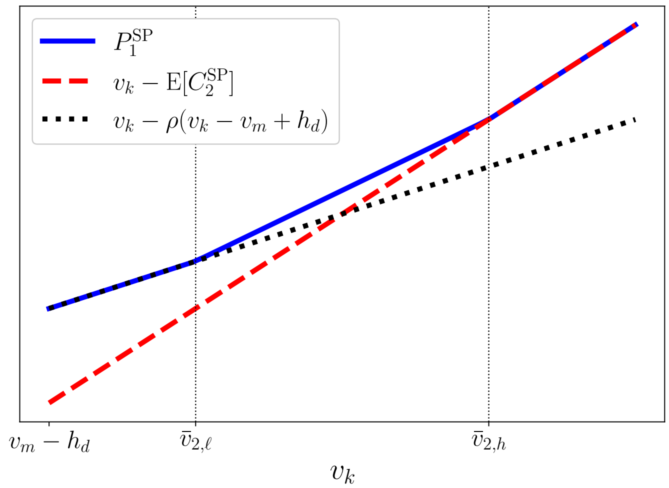
$$P_1^{\mathrm{SP}}(v_k) = v_k - \rho \times \underbrace{\begin{cases} \mathrm{E}[C_2^{\mathrm{SP}}] & \textit{if } v_k > \bar{v}_{2,h}, \\ \pi_h(v_k - v_m + h_d) + (1-\pi_h)c_{2,\ell} & \textit{if } \bar{v}_{2,\ell} \le v_k \le \bar{v}_{2,h}, \\ v_k - v_m + h_d & \textit{if } v_k < \bar{v}_{2,\ell} \end{cases}}_{\textit{expected effective selling cost}} \tag{7}$$

for an MBS of value v_k.

Figure 1 illustrates the impact of having a TBA market at time 2 on time-1 SP prices. Without the TBA market at time 2, any MBS could be sold only on the SP market at time 2, so $P_1^{\mathrm{SP}}(v_k)$ would equal $v_k - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$ for all v_k (the red dashed line). For an MBS whose value $v_k \le \bar{v}_{2,h}$, the existence of the TBA market lowers the expected cost of selling the MBS at time 2 and thus raises the MBS's price in the SP market at time 1 (the blue solid line).

We now describe the equilibrium at time 1. Knowing $P_1^{\mathrm{SP}}(v_k)$, MBS sellers choose between the SP market and the TBA market. If the seller of an MBS with value v_k chooses the SP market, she realizes a net revenue of $P_1^{\mathrm{SP}}(v_k) - C_1^{\mathrm{SP}}$; if the seller chooses the TBA market, she receives P_1^{TBA}. Hence the seller chooses the SP market if $P_1^{\mathrm{SP}}(v_k) - C_1^{\mathrm{SP}} > P_1^{\mathrm{TBA}}$ and the TBA market otherwise. Naturally, the time-1 TBA threshold $\bar{v}_1$ will be the MBS value that equates the revenues from the two markets. Thus, the time-1 equilibrium is as follows.

Figure 1. Time-1 SP price P_1^{SP} as a function of MBS value v_k.



Note: This figure plots how time-1 SP price P_1^{SP} depends on the value of an MBS v_k based on Lemma 1 when $\rho = \pi_h = 0.5$, $c_{2,\ell} = 1$, $c_{2,h} = 3$, $v_m = 10$, $h_d = 5$, and $v_k \in [5,9]$. The red dashed line represents SP buyers' willingness-to-pay at time 1 if they can resell MBSs at time 2 *only* on the SP market; the black dotted line represents SP buyers' willingness-to-pay at time 1 if they can resell MBSs at time 2 *only* on the TBA market; the blue solid line plots P_1^{SP}, which equals SP buyers' willingness-to-pay at time 1 if they can choose between the SP market and the TBA market to resell at time 2 depending on the realization of the SP transaction cost C_2^{SP}.

PROPOSITION 2 (Time-1 equilibrium): *At time* 1*, an MBS with value v_k is sold in the TBA market at price $v_m - h_d$ if $v_k < \bar{v}_1$ and in the SP market at price $P_1^{\mathrm{SP}}(v_k)$ (given by* (7)*) if $v_k \geq \bar{v}_1$ where*

$$\bar{v}_1 := v_m - h_d + \begin{cases} C_1^{\mathrm{SP}} + \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] & \text{if } C_1^{\mathrm{SP}} > c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] \\ \frac{C_1^{\mathrm{SP}} + \rho(1-\pi_h)c_{2,\ell}}{1-\rho\pi_h} & \text{if } (1-\rho)c_{2,\ell} \leq C_1^{\mathrm{SP}} \leq c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] \\ \frac{C_1^{\mathrm{SP}}}{1-\rho} & \text{if } C_1^{\mathrm{SP}} < (1-\rho)c_{2,\ell}. \end{cases} \quad (8)$$

The time-1 TBA threshold $\bar{v}_1$ in general differs from the time-2 TBA threshold $\bar{v}_2$. Depending on parameter values, $\bar{v}_2$ may exceed $\bar{v}_1$. In this situation, because some MBSs sold on the SP market at time 1 may be resold on the TBA market at time 2, the time-1 prices of these SP MBSs depend on the time-2 TBA MBS price, which is lower when MBS heterogeneity h_d is greater. The following result describes the conditions for this situation to occur.

COROLLARY 1 (Impact of parameter values): *If* $C_1^{\mathrm{SP}} > c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$, *MBS heterogeneity* h_d *does not impact the time-*1 *price of any SP MBS. If* $C_1^{\mathrm{SP}} \le c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$, *then* $\bar{v}_1 \le \bar{v}_{2,h}$ *and the time-*1 *prices of SP MBSs whose values fall in* $[\bar{v}_1, \bar{v}_{2,h}]$ *decrease with* h_d.

Intuitively, time-1 SP pricing is completely unaffected by future TBA trading only if MBSs sold on the SP market at time 1 would never be resold on the TBA market at time 2. This requires time-1 SP cost C_1^{SP} to be so high that the time-1 TBA threshold $\bar{v}_1$ exceeds even the highest possible time-2 TBA threshold $\bar{v}_{2,h}$, which may occur but only rarely.[17]

Overall, the key insight from the model is that, because buyers of SP MBSs may use the TBA market as a backup selling venue in the future, the magnitude of the CTD price discount in the TBA market

[17]In Section IA.2 of the Internet Appendix, we provide empirical evidence, based on estimated SP trading costs, that SP pricing is affected by future TBA trading on *at least* 80% of trading days.

can influence the prices and yields of SP MBSs. The more likely an SP MBS today is to be sold into TBA market in the future, the larger the impact of expected CTD discount has on the price of this SP MBS today.

Before developing testable hypotheses, we provide a few discussions on the model setup.

First, generally speaking, we study the impact of transaction costs on asset yields in the spirit of Amihud and Mendelson (1986). The key innovation of our model is the inclusion of two parallel trading mechanisms, leading to the distinctive effect that mitigating the CTD issue in the TBA market can increase MBS prices in the SP market.

Second, because our main focus is on the economic effects of the TBA/SP parallel trading environment, we assume for simplicity that the explicit transaction costs of the two markets and their differences are exogenous, like Amihud and Mendelson (1986). Our main results—on how the CTD issue in the TBA market affects SP yields—would still hold even if the liquidity of TBA and SP markets is endogenous, as long as TBA trading is more liquid than SP trading. Of course, endogenizing TBA market liquidity may deliver further predictions on how MBS heterogeneity affects the TBA liquidity itself, differing from our main focus on the interaction between TBA and SP markets.[18]

Third, two related studies, An et al. (2020) and Huh and Kim (2020), examine how MBS *issuers* take into account the parallel trading envi-

[18]See Li and Song (2020) for a search-based theoretical model along this direction.

ronment when they securitize loans into MBSs, thereby affecting the distribution of MBSs. Our paper takes the distribution of MBSs as exogenously given for two reasons. First, the focus of our paper is on how MBS buyers take into account the potential future selling costs when they bid, thereby influencing MBS prices. Second, our empirical analyses examine the cross-sectional impact of MBS heterogeneity across coupon cohorts that include seasoned coupon cohorts. Variations in MBS heterogeneity for seasoned coupon cohorts are mainly driven by borrowers' refinancing activities, which are exogenous for MBS traders.

Fourth, by assuming that the TBA price equals the value of the cheapest MBS, we shut down a feedback effect from the SP market to the TBA market that could further strengthen the link between current SP prices and future TBA trading. Specifically, suppose that P^{TBA} reflects the average, rather than the lowest, value of TBA MBSs. Then, when a high SP cost $c_{2,h}$ is realized at time 2, MBSs with higher values would be sold into the TBA market, which in turn raises P_2^{TBA}. In consequence, the TBA market at time 2 becomes even more attractive as a backup selling venue for MBS buyers at time 1, thereby enlarging the set of time-1 SP MBSs whose prices depend on the expected MBS heterogeneity at time 2.

B. Testable Hypotheses and Empirical Design

We develop empirically testable hypotheses concerning the impacts of MBS heterogeneity on MBS pricing and trading based on the model presented in Section II.A. We conduct comparative statics with varying levels of the MBS heterogeneity h_d, given a fixed v_m.

When h_d is greater, TBA sellers can deliver worse MBSs and TBA buyers lower their bid prices accordingly, resulting in deeper price discounts $v_k - P_t^{\text{TBA}} = v_k - v_m + h_d$ for MBSs traded on the TBA market. Such CTD discounts are specific to the TBA market (and in fact, are present in all contracts with CTD features, e.g. Treasury futures) and do not depend on the existence of the parallel TBA and SP trading.

In contrast, the dependence of the SP price $P_1^{\text{SP}}(v_k)$ on MBS heterogeneity h_d does reflect the impact of the parallel trading environment. Specifically, the yield of an MBS sold on the SP market at time 1 equals

$$y_1^{\text{SP}}(v_k) := \frac{v_k}{P_1^{\text{SP}}(v_k)} - 1. \tag{9}$$

To see the direction of the impact of h_d more clearly, we examine its impact on $\frac{y_1^{\text{SP}}(v_k)}{1+y_1^{\text{SP}}(v_k)}$, a monotonic transformation of $y_1^{\text{SP}}(v_k)$ that is easier to analyze. Lemma 1 implies that the marginal impact of h_d equals

$$\frac{\partial}{\partial h_d}\left(\frac{y_1^{\mathrm{SP}}(v_k)}{1+y_1^{\mathrm{SP}}(v_k)}\right)=\frac{\rho}{v_k}\times\underbrace{\begin{cases}0 & \text{if } v_k>\bar{v}_{2,h},\\ \pi_h & \text{if } \bar{v}_{2,\ell}\le v_k\le\bar{v}_{2,h},\\ 1 & \text{if } \bar{v}_1\le v_k\le\bar{v}_{2,\ell},\end{cases}}_{\substack{\text{probability of being}\\ \text{resold on the TBA market}}} \tag{10}$$

which is non-negative and decreases with the value of the MBS v_k.

Although SP trading does not involve any CTD issue by design, the yields of SP MBSs whose values fall in the range $v_k \in [\bar{v}_1, \bar{v}_{2,h}]$ do increase with MBS heterogeneity h_d. Intuitively, when the MBS cohort is more heterogeneous, the TBA price falls (relative to the median value of MBSs v_m), which diminishes the value of the future TBA market as a backup selling venue for these SP MBSs. Consequently, buyers lower their bid prices for these SP MBSs to compensate for the drop in potential resale value.

Further, (10) shows that the positive effect of MBS heterogeneity on SP yields is weaker for more valuable MBSs. Intuitively, because more valuable MBSs are less likely to be sold on the TBA market, their yields are less sensitive to the CTD discount in the TBA market resulting from MBS heterogeneity.

We formulate these results as the first testable hypothesis as follows.

HYPOTHESIS 1: *When MBS heterogeneity h_d is greater, the yield of an*

MBS traded on the SP market $y_1^{\mathrm{SP}}(v_k)$ *is higher, and this effect is weaker for a more valuable SP MBS.*

Our second hypothesis concerns the effects of selling pressure, which in the model is captured by ρ, the probability of forced liquidation at time 2. When ρ is greater, (10) shows that the dependence of the yield y_1^{SP} on MBS heterogeneity h_d is stronger. Intuitively, TBA trading as a backup selling venue is more important when SP buyers are more likely to experience liquidity shocks at time 2. We formulate this effect as follows.

HYPOTHESIS 2: *When MBS investors expect heavier selling pressure, the dependence of SP yields on MBS heterogeneity is stronger.*

Our third set of hypotheses concern trading activities on the TBA and SP markets. First, as Proposition 2 shows, a greater MBS heterogeneity h_d results in a lower TBA threshold $\bar{v}_1$. Intuitively, when MBSs are more heterogeneous, TBA buyers expect to receive worse MBSs and lower their bids, thereby raising CTD price discount for any MBS and pushing marginal sellers to the SP market. We state the hypothesis as follows.

HYPOTHESIS 3.1: *When MBS heterogeneity is greater, the proportion of MBSs traded on the SP market is larger.*

Moreover, a greater MBS heterogeneity h_d should dampen trading activities across both the TBA and the SP markets because it raises the trading costs in both markets: CTD discounts in the TBA market are more severe and the cost-saving benefit of TBA trading for SP MBSs diminishes.[19] We formulate this hypothesis as follows.

HYPOTHESIS 3.2: *When MBS heterogeneity is greater, trading is less active on both the TBA and the SP markets.*

Finally, we discuss the impacts of MBS heterogeneity on TBA yields and some related empirical issues when testing the hypotheses.

First, we focus on MBS heterogeneity's effects on yields of SP MBSs rather than that of TBA MBSs. The reason is that the former effect is tied to the parallel trading environment, whereas the latter effect reflects CTD discount that is present even without parallel trading.

Second, we control for a "composition effect" that could result in correlation between the *average* SP prices and MBS heterogeneity, which differs from the economic effect we focus on. An increase in h_d, for example, can simply result from the issuance of MBSs that are worse than the previously cheapest MBS and lead to lower TBA prices. Such lower TBA prices would prompt some sellers to switch from the TBA market to the SP market, reducing the *average* value of SP MBSs. The resulting dependence of *average* SP prices on MBS heterogeneity differs from the

[19]This effect could be incorporated into the model by introducing an explicit MBS holding cost and allowing MBS investors to optimally choose to sell or hold them (a type of market participation cost, as in Vayanos and Wang (2013)). To avoid unnecessary complications, we do not model this channel formally.

liquidity channel we focus on, which affects the yield of every *individual* SP MBS. We shall control for this composition effect by examining SP MBSs with certain fixed characteristics (effectively holding v_k in (10) fixed). See Section III for details.[20]

Third, the effect of MBS heterogeneity on SP pricing arises from a dynamic effect: the pricing of an SP MBS today depends on the *expected* CTD discounts of TBA trading in the future. Hence, the main pricing effect we test is how SP MBS yield $y_t^{\mathrm{SP}}(v_k)$ at time t depends on $\mathrm{E}_t[h_{d,t+1}]$, the time-t expectation of future MBS heterogeneity at time $t+1$. Nonetheless, when taking the model prediction to empirical testing, we focus on the cross-sectional variations of the expected future MBS heterogeneity across coupon cohorts, which is consistent with our hypotheses developed using comparative statics. This cross-sectional analysis helps to exclude confounding effects over time series.

III. Data and Measurement

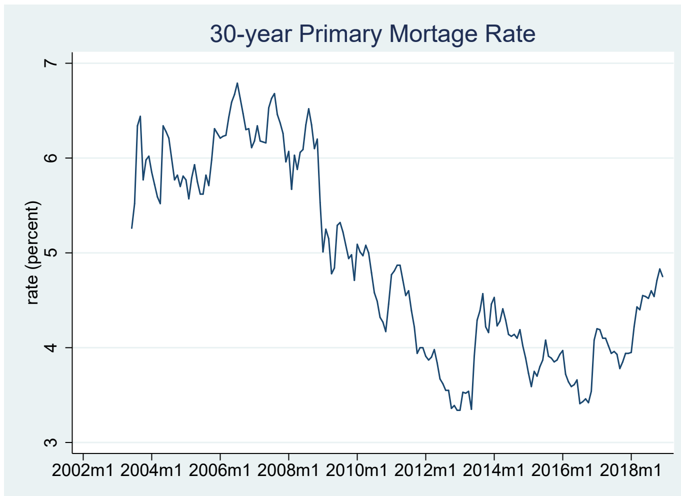
In this section, we introduce the main data sets and measures used in our empirical analyses.

[20]The composition effect could also affect SP yields because worse MBSs usually command higher prepayment risk premiums. Holding v_k fixed teases out this effect. Moreover, for TBA MBSs, the composition effect is empirically challenging to control because we cannot hold the value of the cheapest MBS $v_m - h_d$ constant when MBS heterogeneity h_d changes.

Sample of individual MBSs. Our individual-MBS sample, which is used to compute MBS heterogeneity measures, covers Fannie Mae 30-year MBS coupon cohorts of 2.5%-7% from June 2003 through December 2018. To ensure that we use actively traded cohorts, we limit the sample to coupon cohorts with moneyness in the $[-1.5\%, 4\%]$ range, where the moneyness of a cohort is defined as the difference between the cohort's coupon rate and the current-coupon rate for a synthetic par TBA contract that is obtained by interpolation of TBA prices trading near par.

For each coupon cohort in each month, we obtain prepayment characteristics for each outstanding standard TBA-eligible MBS that belongs to the cohort (excluding Mega securities, stripped MBSs, and collateralized mortgage obligations that are backed by existing MBS, i.e. pools of pools), including the weighted average original FICO score (WAOCS), the weighted average original loan-to-value ratio (WAOLTV), the weighted average original loan size (WAOSIZE), the remaining principal balance (RPB), and the percentage of refinance loans from eMBS through the portal provided by Recursion Co. In constructing heterogeneity measures, we first exclude the set of MBSs that are least likely to be delivered into TBA contracts—based on characteristics following industry practice as described in Himmelberg et al. (2013) and used in Song and Zhu (2019)—and then exclude cohorts with fewer than 1,000 remaining MBSs to ensure that we have sufficiently many MBSs to mea-

Figure 2. Time Series of the Primary Mortgage Rate



Note: This figure plots monthly time series of the 30-year primary mortgage rate (in percentages) from the Freddie Mac survey. The sample period runs from June 2003 through December 2018.

sure cross-sectional heterogeneity. Details of these MBS characteristics and the procedure are provided in Section IA.1 of the Internet Appendix.

In Panel A of Table I we present summary statistics for the sample period and moneyness for each included coupon cohort. Overall, the sample comprises an unbalanced panel, with the general sample period running from June 2003 through December 2018 but with varying starting months for various cohorts. Given the downward trend in mortgage rates in the sample period (as shown in Figure 2), higher coupon cohorts appear in the earlier part and lower coupon cohorts appear in the later part of the sample. The time-series mean of moneyness, which ranges between −0.82% and 2.46%, is increasing in the cohort coupon rate.

In Panel B of Table I we report summary statistics for the number of MBSs for each included coupon cohort. Specifically, for each cohort i in month t, we count the total number of MBSs N_{it}. Then, for each cohort i, we report the minimum, quartiles, and maximum of the monthly series N_{it}. The median number of MBSs is the largest for the 5.5% and 6% cohorts, and is smaller for cohorts with lower and higher coupons. This is because mortgage rates only reached very low and high levels in short periods of time in our sample, as Figure 2 shows. The minimum number of MBSs is around 1,000 for cohorts of coupons 2.5%-5.5% but about 7,000-16,000 for cohorts of coupons 5.5%-6.5%. The 25th percentiles are over 4,900 for most coupon cohorts. Overall, the number of MBSs within each cohort is sufficient to measure heterogeneity.

MBS prepayment characteristics. As discussed in Section I, prepayment is the most important determinant of MBS value. To capture the heterogeneity of MBS values, we use WAOCS, which is a key input for prepayment models (Fabozzi and Mann (2011) and Hayre (2001)). An appealing feature of WAOCS is that a high WAOCS is usually associated with high prepayment risk and low MBS value.[21] We also obtain

[21]In Section IA.1 of the Internet Appendix, we analyze the effects of various prepayment characteristics, including WAOCS, WAOLTV, and WAOSIZE, on prepayment rates using individual-MBS-level regressions. We also conduct robustness checks us-

Table I. Summary Statistics for Monthly CUSIP-Level MBS Characteristics

A: Sample and Moneyness

	Sample			Moneyness			
Coupon	Begin	End	N	mean	sd	min	max
2.5	2017/04	2018/12	20	-0.82	0.35	-1.50	-0.35
3	2012/08	2018/12	77	-0.02	0.46	-1.08	0.89
3.5	2011/04	2018/12	93	0.44	0.48	-0.78	1.39
4	2009/06	2018/12	115	0.73	0.64	-0.92	1.89
4.5	2003/10	2018/12	175	0.50	1.17	-1.48	2.39
5	2003/06	2018/12	187	0.89	1.23	-1.38	2.89
5.5	2003/06	2018/12	187	1.39	1.23	-0.88	3.39
6	2003/06	2018/12	187	1.89	1.23	-0.38	3.89
6.5	2003/06	2018/12	174	2.25	1.17	0.12	4.00
7	2003/06	2018/12	145	2.46	1.05	0.62	4.00

B: Summary Statistics for the Number of CUSIPs

Coupon	min	p25	p50	p75	max
2.5	1001	1003	1008	1011	1014
3	1113	8767	11049	15640	16006
3.5	1004	10598	18106	28652	33710
4	1097	7331	17196	27159	35220
4.5	1029	2513	15509	20204	23633
5	1481	14380	20006	22730	24859
5.5	8883	25108	29581	35075	37314
6	16537	22545	26960	34527	38801
6.5	6973	12356	22981	24970	29916
7	1955	4919	9529	10650	18052

C: Time Series Means of Cross-Sectional Percentiles of WAOCS and SMM

	WAOCS					SMM				
Coupon	p5	p25	p50	p75	p95	p5	p25	p50	p75	p95
2.5	747	768	775	781	792	0.00	0.17	0.45	1.30	25.50
3	744	760	767	773	783	0.01	0.17	0.44	4.43	29.83
3.5	722	748	760	769	781	0.00	0.14	0.41	5.90	39.78
4	716	741	755	766	779	0.01	0.15	0.49	10.53	43.41
4.5	707	731	746	757	772	0.00	0.14	0.61	10.81	45.17
5	699	719	731	743	763	0.01	0.16	1.24	14.91	52.37
5.5	691	710	722	735	758	0.01	0.15	2.11	18.02	60.11
6	687	703	716	731	758	0.00	0.30	2.69	18.13	67.22
6.5	684	698	712	728	758	0.00	0.54	3.02	11.20	68.28
7	683	695	708	725	756	0.00	0.05	2.30	10.24	66.89

Note: Panel A reports a summary of the included coupon cohorts, including the beginning month, the ending month, the number of monthly observations (N) as well as the time-series percentiles of moneyness for each coupon cohort. The moneyness, in percentage, equals the difference between the cohort's coupon rate and the coupon rate for a synthetic par TBA contract interpolated using TBA prices trading near par. Panel B reports, for each coupon cohort, the percentiles of the monthly time series of the number of outstanding MBSs. Panel C reports the means of the monthly time-series of the percentiles of WAOCS and SMM within a coupon cohort. The overall sample period runs from June 2003 through December 2018, and includes FNMA 30-year TBA-eligible MBSs.

the realized prepayment rate for each MBS within each coupon cohort for each month, known as the single monthly mortality rate (SMM), which equals the fraction of the scheduled balance (= total beginning balance − scheduled principal payment) at the beginning of the month that was prepaid during that month.[22]

Panel C of Table I presents time-series means of the percentiles of WAOCS and SMM for each coupon cohort. In particular, for each MBS j in cohort i in month t, we observe the WAOCS_{itj} and SMM_{itj}. We compute the 5th, 25th, 50th, 75th, and 95th percentiles of WAOCS_{itj} and SMM_{itj} across MBS $j = 1, \cdots, N_{it}$ for each cohort i in month t. We then compute the time series average of these five percentiles, for each cohort i.

We observe that all the percentiles of WAOCS show a sharply decreasing pattern in the cohort coupon rate, indicating a shift in the distribution to the high-WAOCS region when the mortgage rate decreases. This pattern arises because in MBSs issued earlier in the sample with high coupon rates, high FICO loans refinanced more quickly and dropped out of the MBS when the mortgage rate decreased, after which the refinanced loans are then packaged into new MBS with lower

ing heterogeneity measures based on WAOSIZE, a combination of different characteristics, and prepayment rates. See Section IA.4 of the Internet Appendix for details.

[22]The SMM can be converted into the annualized constant prepayment rate (CPR) by $\text{CPR} = 1 - (1 - \text{SMM})^{12}$.

coupon rates. That is, the high prepayment speed associated with high FICO scores, together with the decreasing trend in the mortgage rate, leads to the rightward shift in the distribution of WAOCS (across MBSs within a cohort) from high to low coupon cohorts. We also observe that the percentiles of SMMs generally increase with cohort coupons, confirming the higher prepayment speeds of deeper in-the-money cohorts. The lower SMM of the 7%-cohort when compared with the slightly lower coupon cohorts is consistent with a burnout effect (Hayre (2001)).

MBS yields and returns. We follow relevant studies, such as Gabaix et al. (2007), Boyarchenko et al. (2019), and Song and Zhu (2019), to use the OAS in our empirical analyses. The OAS is the interest rate spread added to the term structure of interest rates such that the present value of the expected future cash flows of an MBS, after adjusting for the value of homeowners' prepayment options, equals the market price of the security. We obtain the OAS series based on the Treasury term structure of FNMA 30-year SP MBSs over June 2012-December 2018 from a major Wall Street MBS dealer.[23]

Specifically, for each coupon cohort in each month, we obtain the month-end OAS for six groups of SPs with LTV below 90%, from 90%

[23]There are several potential issues with OAS measures, such as prepayment model misspecifications, non-interest-rate prepayment risk premiums, and so on. We address these issues in Section IV.E and Section IA.4 of the Internet Appendix.

to 95%, from 95% to 100%, from 100% to 105%, from 105% to 125%, and above 125%. With 105% as the threshold, the first four groups are eligible for TBA trading and more valuable than the last two groups ineligible for TBA trading. Among the TBA-eligible MBSs, higher-LTV groups usually have lower prepayment risk and are of higher value.[24] Using the SPs with fixed characteristics is important because it controls for the composition effect as discussed in Section II.B.

We match OAS series to the MBS characteristics sample and exclude those without a match. Panel A of Table II provides a summary of the SP OAS sample. Specifically, the series start in June 2012 for the 3.5%-4.5% coupon cohort, in July 2012 for the 5% coupon cohort, and in October 2012 for the 3% coupon cohort. The time series average of the number of outstanding MBSs is more than 10,000 for all coupon cohorts. Panel B reports the time-series means of the SP OAS for all available coupon cohorts. We observe that the mean OAS is higher for higher coupon cohorts that are deeper in the money, except that the OAS of the 5% cohort is lower than those of the lower coupon cohorts. More-

[24] Consistently, based on the IHS Markit Agency RMBS Specified Pool Summary of December 2016, the payups are *higher* for SPs with higher LTV ratios in general, but the payups on SPs with LTV ratios higher than 105% are slightly *lower* than those with LTV ratios between 100% and 105%. Details are available at https://cdn.ihs.com/www/blog/commentary/pdf/Markit-Agency-RMBS--Specified-Pool-Summary--December-2016.pdf.

over, within each coupon cohort, the mean OAS is higher for SPs with higher LTV ratios. This is consistent with our model's implications that because low-LTV MBSs benefit more from the existence of TBA trading, they enjoy a higher liquidity premium, which results in lower yields.

We also obtain the OAS series for TBA contracts based on the Treasury term structure for FNMA 30-year MBSs with coupon rates ranging from 2.5% to 7% over June 2003-December 2018 from the same MBS dealer.[25] Panel C of Table II reports the summary statistics the TBA OAS sample. In terms of cohort×month, the TBA OAS sample is the same as the MBS characteristics sample. The mean OAS is also higher for higher coupon cohorts that are deeper in the money.

Transaction data. To measure MBS trading activities, we use the TRACE dataset of MBS transactions that the FINRA began collecting in May 2011. Each trade record contains the trade type, agency, loan terms, security coupon rate, price, par value, trade date, and settlement month among other features for each trade. Both inter-dealer trades and trades between dealers and customers are included.

For TBA trades, we keep the regular good delivery outright transac-

[25] In constructing monthly series of the TBA OAS, we use the value on the last business day of the first week in a month, which is among the days with the most active trading activity (Gao et al. (2017)). Further, we use the OAS for the front-month TBA contracts, which usually settle in the second week of the same month.

Table II. Summary Statistics of Monthly OAS Series

A: SP Sample						
Coupon	Begin	End	N	Average Moneyness	Average # CUSIP	
3	2012/10	2018/12	75	-0.03	11518	
3.5	2012/06	2018/12	79	0.50	21601	
4	2012/06	2018/12	79	1.00	23359	
4.5	2012/06	2018/12	79	1.50	19952	
5	2012/07	2018/12	78	1.99	21782	
B: Time Series Means of SP OAS						
Coupon	80-90	90-95	95-100	100-105	105-125	> 125
3	27.14	29.66	33.24	34.15	40.50	50.59
3.5	29.50	34.42	36.17	39.76	41.87	50.11
4	38.96	40.67	41.52	41.93	42.48	47.54
4.5	41.26	36.52	40.25	37.12	48.78	50.62
5	27.08	27.32	26.61	26.24	26.59	48.97
C: TBA Sample						
Coupon	Begin	End	N	Average Moneyness	Average # CUSIP	OAS
2.5	2014/04	2018/12	20	-0.82	1007	20.36
3	2012/08	2018/12	77	-0.02	11254	16.06
3.5	2011/04	2018/12	93	0.44	18639	16.68
4	2009/06	2018/12	115	0.73	17523	18.82
4.5	2003/10	2018/12	175	0.50	12152	37.96
5	2003/06	2018/12	187	0.89	18219	32.57
5.5	2003/06	2018/12	187	1.39	28855	35.05
6	2003/06	2018/12	187	1.89	28126	37.44
6.5	2003/06	2018/12	174	2.25	19268	62.94
7	2003/06	2018/12	145	2.46	8581	86.34

Note: Panel A reports a summary of the FNMA 30-year SP OAS sample, including the beginning month, the ending month, and the number of monthly observations as well as the mean of the monthly time-series the moneyness and the number of all outstanding MBS within each cohort. Panel B reports the mean of the monthly OAS series, for each coupon cohort of each group of SP MBSs. Panel C reports the summary of the TBA OAS series. The overall sample period runs from June 2012 through December 2018 for SP, while from June 2003 through December 2018 for TBA.

tions of FNMA 30-year MBSs in the standard coupon cohorts of 2.5%-7%.[26] In matching SP trading activities, we only use trades of front-month TBA contracts. For each coupon cohort, we compute both the total par dollar trading volume and the number of trades of front-month TBA contracts in each month. This usually spans a period running from the day after the TBA settlement day in the previous month to the settlement day in the current month. For SP trades, we keep the transactions of FNMA 30-year TBA-eligible pass-through securities with the same standard coupons of 2.5%-7% as TBA trades. Similar to the aggregation of TBA trades, for each coupon cohort we compute the total par dollar trading volume and number of trades of SP MBSs from the day after the TBA settlement day in the previous month to the settlement day in the current month.

We keep only the cohort×month for which both TBA and SP trading activity measures are available. We then match the transaction data to the MBS characteristics data and exclude those without a match. In Panel A of Table III, we report the sample summary. The sample period runs from June 2011 through July 2015 for each of the 3.5%-6.5% coupon cohorts. Yet, the number of observations varies because trading activity measures are not always available during the period.

[26]Trades involving stipulated TBA contracts and dollar rolls, as well as those not qualified for good delivery and with quarter or non-standard coupon rates, are hence excluded.

Table III. Summary of Monthly TBA and SP Trading Activity

A: Sample

Coupon	Begin	End	N	Average Moneyness	Average # CUSIP	Average Outstanding	Average Issuance
3	201208	201507	36	0.03	7676	291.634	11.164
3.5	201106	201507	50	0.48	10777	273.778	11.863
4	201106	201507	50	0.98	14676	334.270	8.697
4.5	201106	201507	50	1.48	17926	295.472	2.223
5	201106	201507	50	1.98	23585	197.462	0.381
5.5	201106	201507	50	2.48	32854	165.175	0.025
6	201106	201507	50	2.98	28873	107.067	0.017
6.5	201106	201507	41	3.34	15634	35.992	0.007

B: Monthly Average Activity of All Trades

	Dollar Volume ($billion)			Number of Trades		
Coupon	TBA	SP	SP/TBA	TBA	SP	SP/TBA
3	221.08	14.89	0.07	11039	1886	0.23
3.5	283.77	15.91	0.05	13388	2609	0.20
4	246.31	19.82	0.08	10926	3334	0.37
4.5	124.85	14.01	0.15	5251	3093	0.79
5	54.93	5.92	0.40	2284	1454	1.07
5.5	25.88	4.48	0.98	1216	1572	2.69
6	12.24	3.17	2.68	699	1183	5.34
6.5	1.03	0.75	9.40	110	449	23.59

C: Monthly Average Activity of Dealer-Customer Trades

	Dollar Volume ($billion)			Number of Trades		
Coupon	TBA	SP	SP/TBA	TBA	SP	SP/TBA
3	110.58	13.73	0.13	2459	1398	0.89
3.5	129.18	14.54	0.11	2858	1803	0.67
4	109.38	17.41	0.17	2329	2405	1.18
4.5	52.70	11.76	0.26	1032	2177	2.53
5	22.00	5.09	0.62	436	1003	2.72
5.5	10.68	3.85	1.53	231	1071	5.82
6	5.65	2.72	3.08	143	813	8.97
6.5	0.41	0.46	17.46	32	267	25.72

Note: In Panel A we report summary statistics for the sample of monthly TBA and SP trading activities of FNMA 30-year MBS, including the beginning month, the ending month, and the number of monthly observations as well as the means of the monthly time series of moneyness, the number of all outstanding MBS, the total outstanding balance (in $billion), and total new issuance (in $billion), for each coupon cohort. Panel B reports the means of the monthly time-series of the SP and TBA trading activity measures and their ratios, in both $billlion volume and the number of trades using all trades. Panel C reports similar summary statistics but using only dealer-customer trades. We consider front-month TBA contracts and aggregate the SP trades of standard pass-through securities for a period running from the day after the TBA settlement day in the previous month to the settlement day in the current month. The overall sample runs from June 2011 through July 2015 based on TRACE data of agency MBS transactions.

The sample has a shorter time period for the 3% coupon cohort, running from August 2012 through July 2015. The average moneyness is all positive, increasing with coupon rate from 0.03 to 3.34, whereas the average number of outstanding MBSs within a coupon cohort is larger than 10,000 for all except the 3% cohort.

The last two columns report the time-series average of the total outstanding balance and new issuance (both in $billions) for each coupon cohort, obtained from eMBS. The outstanding balance is higher than $100 billion for all except the 6.5% cohort. It decreases from low to high coupons because of the low levels and decreasing trend of mortgage rates during the sample period of June 2011-July 2015. The average monthly new issuance also decreases from low to high coupons: the issuance is more than $2 billion a month for 3%-4.5% but less than $0.5 billion a month for coupons higher than 4.5%. The high outstanding balance but low new issuance of 5%-6% coupon cohorts occurs because these cohorts experienced active issuance in periods leading to June 2011.

In Panels B and C of Table III we report the means of monthly time-series of the SP and TBA trading activities and their ratios, measured with both dollar volume and number of trades. Panel B includes both inter-dealer and dealer-customer transactions, while Panel C includes only dealer-customer transactions. We observe that both SP trading and TBA trading are more active in low-coupon cohorts. The SP/TBA ratio of trading activity, however, increases monotonically with coupons. This pattern is strong whether all trades or only dealer-customer trades are included and whether dollar volume or number of trades is used.

Time-series variables. We construct the balance-sheet-based leverage ratio measure of the aggregate intermediary sector proposed by He et al. (2017), and calculate the market-price-based "noise" measure proposed in Hu et al. (2013). The leverage-ratio measure is computed as the aggregate market equity plus aggregate book debt divided by aggregate market equity, using CRSP/Compustat and Datastream data, of the holding companies of primary dealers recognized by the FRBNY. The "noise" measure is computed as the root mean squared distance between the market yields of Treasury securities and the hypothetical yields implied from yield curve models like that of Svensson (1994).[27] Both variables are available at daily frequency for our sample period; we use their values on the last business day of the first week in each month to construct monthly series, in a manner similar to the construction of monthly TBA OAS series discussed above. Moreover, we follow He et al. (2019) to use the first principal component of the leverage ratio

[27] The Svensson (1994) model is used to construct Treasury yield curves that are regular inputs in the Federal Reserve's policy discussions and publications (Gurkaynak et al. (2007)), and also used by the Federal Reserve in evaluating offers submitted in auctions through which the purchases of Treasury securities for quantitative easing are executed (Song and Zhu (2018)).

and "noise" as a parsimonious measure of financial intermediary constraints (and ρ in our model), denoted as *Distress*.

In addition, the mortgage rates used in Figure 2 are the 30-year fixed-rate mortgage loan rates from the Freddie Mac Primary Mortgage Market Survey (PMMS), available at weekly frequency. We use the value of PMMS in the first week of each month to construct the monthly series.

IV. Economic Effects of MBS Heterogeneity

In this section, we empirically test the impacts of MBS heterogeneity. To be clear, our tests examine variations in MBS heterogeneity *across* coupon cohorts, with the associated hypotheses in Section II.B developed based on comparative statics. Nevertheless, MBS heterogeneity influences MBS yields through a dynamic channel. In particular, the SP MBS yield $y_t^{\mathrm{SP}}(v_k)$ at time t depends on $\mathrm{E}_t[h_{d,t+1}]$, the time-t expectation of future MBS heterogeneity at time $t+1$. Hence, we start with introducing the measures of MBS heterogeneity and examine their time-series features.

A. *Measures of MBS Heterogeneity*

We empirically measure MBS heterogeneity—the value of the cheapest MBS relative to the cohort median (h_d as defined in Section II.B)—using the prepayment characteristic of the cheapest MBS relative to the average characteristic of all MBSs within a coupon cohort. In particular, we define

$$h_{it}^{\mathrm{WAOCS}} = \mathrm{WAOCS}_{it}^{95\%} - \mathrm{WAOCS}_{it}^{50\%}, \tag{11}$$

where $\mathrm{WAOCS}_{it}^{95\%}$ and $\mathrm{WAOCS}_{it}^{50\%}$ are the 95th percentile and median, respectively, of the WAOCS across all N_{it} MBSs within coupon cohort i in month t. Given that MBS value monotonically decreases with WAOCS, h_{it}^{WAOCS} captures the value of the cheapest MBS relative to the average MBS. We use the 95th percentile rather than the maximum to avoid the impact of outliers.

We empirically proxy the time-t expectation of future MBS heterogeneity by h_{it}^{WAOCS}. To investigate whether this measure performs well, we construct a measure of heterogeneity using realized prepayment rates directly h_{it}^{SMM} for coupon cohort i at month t as follows:

$$h_{it}^{\mathrm{SMM}} = (\mathrm{SMM}_{it}^{95\%} - \mathrm{SMM}_{it}^{50\%}) \times \mathrm{ITM}_{it} + (\mathrm{SMM}_{it}^{50\%} - \mathrm{SMM}_{it}^{5\%}) \times \mathrm{OTM}_{it}, \tag{12}$$

where ITM_{it} and OTM_{it} are indicator variables for whether the coupon cohort i is in-the-money or out-of-the-money at month t. We use the

95th percentile for in-the-money cohorts and the 5th percentile for out-of-the-money cohorts because prepayment hurts premium MBSs but benefits discount MBSs.[28] We provide summary statistics of h_{it}^{WAOCS} and h_{it}^{SMM} in Table IA.II of the Internet Appendix.

To verify whether h_{it}^{WAOCS} captures the heterogeneity of future prepayment rates well, we consider the following regression:

$$h_{i,t+n}^{\text{SMM}} = \beta \cdot h_{it}^{\text{WAOCS}} + \text{FE}_{\text{Moneyness}} + \varepsilon_{it}, \tag{13}$$

where $h_{i,t+n}^{\text{SMM}}$ is the average of the heterogeneity measure of the realized prepayment rate from month $t+1$ to $t+n$ for cohort i. The moneyness fixed effect is included, so the coefficient β captures whether MBS heterogeneity in prepayment rates in future months depends on MBS heterogeneity in WAOCS in the current month for a given moneyness cohort.

In the first three columns of Table IV, we report results of the regression in (13) for n=1, 3, and 12 months, respectively. The regression coefficients β are positive and highly significant for all three horizons and are lower for longer horizons n. In the last three columns we report similar regressions using $h_{i,t+n}^{\text{WAOCS}}$ as the dependent variable. That

[28]A coupon cohort is in-the-money (out-of-the-money) if the moneyness of MBSs within this cohort is positive (negative). Premium (discount) MBSs are MBSs that fall within in-the-money (out-of-the-money) cohorts.

Table IV. MBS Heterogeneity Measures

	h^{SMM}_{t+1}	$h^{\text{SMM}}_{t+1,t+3}$	$h^{\text{SMM}}_{t+1,t+12}$	h^{WAOCS}_{t+1}	$h^{\text{WAOCS}}_{t+1,t+3}$	$h^{\text{WAOCS}}_{t+1,t+12}$
h^{WAOCS}_{t}	0.40**	0.37*	0.20**	0.99***	0.98***	0.52***
	(2.04)	(1.90)	(1.97)	(467.59)	(304.16)	(97.27)
Intercept	-9.97**	-9.38*	-3.61	0.41***	0.73***	1.15***
	(-1.98)	(-1.84)	(-1.37)	(7.13)	(9.36)	(8.30)
Obs	1,521	1,497	1,389	1,521	1,497	1,389
R^2_{adj}	0.74	0.76	0.65	1.00	1.00	0.99
Moneyness FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports panel regressions of $h^{\text{WAOCS}}_{i,t+n}$ and $h^{\text{SMM}}_{i,t+n}$—the time series average of the heterogeneity measures over $t+1$ to $t+n$ for n=1, 3, and 12 months—on $h^{\text{WAOCS}}_{i,t}$, with moneyness-cohort fixed effects included. The overall sample period runs from June 2003 through December 2018. We report t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions in parentheses. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

is, these regressions examine whether MBS heterogeneity in WAOCS in the current month forecasts that in future months. The regression coefficients are also positive and highly significant for all three horizons and are lower for longer horizons n. Overall, the results show that investors can form reasonably accurate expectations of future MBS heterogeneity.

Finally, we briefly discuss the variation in h^{WAOCS} across coupon cohorts. Recall that the FICO score for each loan that is used to compute the WAOCS of an MBS is its original value at issuance, while the loan balance used as the weight is the remaining loan balance. The WAOCS of an MBS may vary over time after issuance because the remaining balances of loans within the MBS may evolve because of prepayment. In consequence, the cross-sectional variation of h^{WAOCS} is driven both by the variation at issuance and the variation that emerges after issuance because of refinancing activities.

B. *MBS Heterogeneity and Yields*

In this section, we empirically test the effect of MBS heterogeneity on SP returns, which is a distinctive economic effect of the parallel trading environment, as formulated in Hypothesis 1. Specifically, we test whether MBS heterogeneity positively affects SP yields and whether this effect is stronger for MBSs that are more likely to be delivered into TBA contracts.

We consider the following panel regression over cohort i and month t:

$$\text{OAS}_{itj} = \beta_{1j} \cdot h_{it}^{\text{WAOCS}} + \beta_{2j} \cdot \text{SMM}_{itj} + \beta_{3j} \cdot \text{WAOLTV}_{it} + \text{FE}_{\text{Time}} + \varepsilon_{itj} \quad (14)$$

for each j, where j represents one of the six types of SPs based on LTV ratios. Time fixed-effects are included, so the coefficient β_{1j} captures the effects of MBS heterogeneity on the cross-sectional variation of OAS. We control for the prepayment rate SMM_{itj}. Moreover, because each group of SPs only fixes a range of LTV ratio, LTV ratios may still vary across the SPs within a LTV group. To control for such potential variations across the dimension of coupon cohort i, we compute the average of WAOLTV of MBSs within the cohort i at month t, weighted by the remaining principal balance. We denoted this measure as WAOLTV_{it} and include it as a control. This is important especially when h_{it}^{WAOCS} is correlated with WAOLTV_{it} because of past refinancing activities.

In Panel A of Table V we report the results of the panel regression in (14) for TBA-eligible SP MBSs with LTV ratios lower than 105% in the first four columns. We observe that h^{WAOCS} significantly affects the OAS positively, consistent with our model's prediction that having future TBA trading as an option affects current SP prices. The effect is weaker for those with higher LTV ratios that are less likely to be delivered into TBA contracts. Moreover, the last two columns report the regression results for TBA-ineligible SP MBSs with LTV ratios higher than 105%. We observe that the regression coefficients on h^{WAOCS} are much lower and statistically insignificant.

The economic magnitudes of the effects of MBS heterogeneity are also large. For example, a one-standard-deviation increase of h^{WAOCS} across coupon cohorts (about 11.81 based on the between standard deviation) is associated with an increase in OAS by about 17 ($\approx 11.81 \times 1.41$) basis points for SP MBSs with LTV ratios in the 80-90% range, and by about 10 ($\approx 11.81 \times 0.83$) basis points for SP MBSs with LTV ratios in the 100%-105% range. That is, the effects diminish by almost half for

Table V. MBS Heterogeneity and Yields

	TBA-Eligible SP (LTV)				TBA Ineligible SP (LTV)		TBA
	80-90	90-95	95-100	100-105	105-125	> 125	
A: Regression on h^{WAOCS}, SMM, and WAOLTV							
h^{WAOCS}	1.41***	1.33***	1.12***	0.83**	0.51	0.31	1.43***
	(5.16)	(4.10)	(3.20)	(2.08)	(1.39)	(0.97)	(4.66)
SMM	-1.62***	-1.97***	-2.11***	-2.10***	-2.36***	-1.65***	-0.71***
	(-5.13)	(-6.03)	(-5.50)	(-5.45)	(-3.54)	(-3.59)	(-3.70)
WAOLTV	-1.21	-0.77	-0.77	0.64	3.02	-0.61	2.68*
	(-1.20)	(-0.69)	(-0.51)	(0.42)	(1.32)	(-0.83)	(1.94)
Intercept	99.26	83.06	92.02	1.21	-149.03	121.66**	-140.85*
	(1.41)	(1.04)	(0.85)	(0.01)	(-0.92)	(2.46)	(-1.65)
Obs	390	390	390	390	390	390	1,360
R^2_{adj}	0.64	0.63	0.56	0.55	0.53	0.60	0.74
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B: Regression on SMM and WAOLTV							
SMM	-0.51**	-0.80***	-1.05***	-1.26***	-1.69**	-1.16***	0.24
	(-2.01)	(-2.88)	(-3.13)	(-3.56)	(-2.28)	(-2.42)	(1.33)
WAOLTV	2.66*	3.37*	2.85	3.31	3.92	-0.21	3.69**
	(1.83)	(1.83)	(1.34)	(1.50)	(1.59)	(-0.30)	(2.30)
Intercept	-158.90	-193.43	-149.83	-176.35	-206.02	97.20**	-205.84**
	(-1.55)	(-1.48)	(-1.00)	(-1.13)	(-1.20)	(1.99)	(-1.98)
Obs	390	390	390	390	390	390	1,360
R^2_{adj}	0.49	0.50	0.47	0.50	0.52	0.59	0.69
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: In this table we report the results for panel regressions of the OASs of four groups of TBA-eligible SP MBSs (first four columns), of two groups of TBA-ineligible SP MBSs (last two columns), and of TBA MBSs, using the sample of FNMA 30-year MBSs. In Panel A we report the results of regressions on h^{WAOCS} controlling for SMM and WAOLTV, while Panel B reports regressions on SMM and WAOLTV. The control variable WAOLTV used in regressions of SP MBSs is the average WAOLTV of MBSs within the cohort i at month t, weighted by the remaining principal balance, while that used in regressions of TBA MBSs is the 5th percentile. Time dummies are included, and t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. The overall sample period runs from June 2003 through December 2018 for TBA MBSs and from June 2012 through December 2018 for SP MBSs. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

SP MBSs that are unlikely to be delivered into TBA contracts.[29]

To check the overall explanatory power of MBS heterogeneity, we report panel regression without h^{WAOCS} in Panel B of Table V. We observe that the increase of adjusted R^2 when including h^{WAOCS} ranges from 5% to 15%.

In addition, although our main focus is the effect of MBS heterogeneity on SP yields, we also run the regression in (14) for TBA MBSs. As reported in the last column in Panel B of Table V, h^{WAOCS} significantly affects OAS of TBA MBSs positively, consistent with the CTD discount.[30] A one-standard-deviation increase of h^{WAOCS} across coupon cohorts is associated with an increase in OAS by about 18 basis points for TBA MBSs, similar to SP MBSs with LTV ratios in the 80-90% range that are very likely to be delivered into TBA contracts.

[29] One may worry that h^{WAOCS} may simply be correlated with SP trading costs and affect MBS returns through its liquidity impact (Amihud and Mendelson (1986)). Nevertheless, Gao et al. (2017) show that while the SP trading cost does decrease with the LTV ratio, it increases substantially across the 105% threshold. Hence the weaker effects of h^{WAOCS} on SP MBSs with LTV ratios higher than 105% are inconsistent with this alternative interpretation.

[30] Measures of TBA yields are usually computed using the TBA price and a set of MBSs that are representative of TBA deliveries that likely contain MBSs more valuable than the cheapest. The CTD discount is naturally included in these TBA yield measures.

C. Liquidity Shocks

We now examine whether the effects of MBS heterogeneity on SP yields are stronger when selling pressure is heavier, i.e. Hypothesis 2.

As discussed in Section III, we use the *Distress* measure to proxy for the probability of liquidity shocks (ρ in our model), which has been shown to capture the extent of investment capital constraints. We consider the following panel regression

$$\begin{aligned}\mathrm{OAS}_{itj} = &\beta_1 \cdot h_{it}^{\mathrm{WAOCS}} + \beta_2 \cdot h_{it}^{\mathrm{WAOCS}} \times \rho_t + \beta_3 \cdot \rho_t + \beta_4 \cdot \mathrm{SMM}_{itj} \\ &+ \beta_5 \cdot \mathrm{WAOLTV}_{it} + \mathrm{FE}_{\mathrm{Time}} + \mathrm{FE}_{\mathrm{SPType}} + \varepsilon_{itj},\end{aligned} \tag{15}$$

for the whole SP sample by pooling all six types of SP MBSs together. We control for SMM_{itj} and WAOLTV_{it} and include time fixed-effects. As we pool all types of SP MBSs together to improve the accuracy of statistical inference, we include a SP-type fixed-effect accordingly.

We observe from column (1) of Table VI that the coefficient on the interaction term $h_{it}^{\mathrm{WAOCS}} \times \rho_t$ is positive and highly significant, confirming that the effects of MBS heterogeneity on SP yields are stronger when selling pressure is higher. In column (2), we report the regression with time fixed-effects, which absorb all time-series variables. The interaction term using *Distress* is still positive and highly significant.

Table VI. Liquidity Shocks

	(1)	(2)
h^{WAOCS}	1.09***	1.33***
	(2.80)	(3.46)
$h^{\text{WAOCS}}\times$Distress	3.13***	4.24***
	(3.48)	(4.07)
Distress	-54.20**	
	(-2.57)	
SMM	-1.64***	-1.61***
	(-4.59)	(-6.74)
WAOLTV	1.27	1.94
	(0.80)	(1.07)
Intercept	-63.11	-96.25
	(-0.58)	(-0.71)
Obs	1,620	1,620
R^2_{adj}	0.43	0.60
Time FE	No	Yes
SP Type FE	Yes	Yes

Note: We report in this table the results for panel regressions of the SP OASs on the interaction terms $h^{\text{WAOCS}}\times$*Distress* using monthly data of FNMA 30-year MBSs. We pool all six groups of SP MBSs, including fixed-effects for SP types. Time fixed-effects are excluded in the regression reported in column (1), where *Distress* is controlled for directly, but are included in the regression reported in column (2). All regressions include SMM and WAOLTV as controls. The t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. The overall sample period runs from June 2012 through December 2018. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

D. *Trading Activities*

Our third set of hypotheses concerns the effects of MBS heterogeneity on MBS trading activities. In this section, we examine whether the ratio of SP to TBA trading activity increases with heterogeneity (Hypothesis 3.1) and whether TBA and SP trading activities both weaken with MBS heterogeneity (Hypothesis 3.2).

In Columns (1)-(2) of Panels A of Table VII, we report regressions of the monthly dollar volume of TBA and SP trading, respectively, on h_{it}^{WAOCS}. In addition to time fixed effects, we include monthly issuance amounts to control for the supply of MBSs.[31] Not surprisingly, we find that issuance positively affects TBA and SP trading activities. Importantly, h_{it}^{WAOCS} significantly affects MBS trading activities after controlling for issuance. Specifically, the regression coefficients on h_{it}^{WAOCS} are significantly negative for both TBA and SP trading volume, confirming that trading activity indeed weakens when MBS heterogeneity is greater. Further, the regression of the log ratio of SP to TBA trading volume, reported in column (3), shows significantly positive coefficients on h_{it}^{WAOCS}. In sum, consistent with our model's predictions, when MBS heterogeneity is greater, a larger proportion of MBSs are sold through the SP market rather than the TBA market because deeper TBA price

[31]The results controlling for outstanding balance are similar, as presented in Section IA.4 of the Internet Appendix.

discounts prompt sellers of more valuable MBSs to prefer SP trading.

The results are similar when we use the total number of trades to measure trading activities, as reported in columns (4)-(6) of Panel A, and when we use dealer-customer trades, as reported in Panel B. An interesting observation is that for TBA trading activity, the magnitudes of the regression coefficients are appreciably lower for dealer-customer trades than for all trades; such a pattern is not present for SP trading activity. This suggests that inter-dealer TBA trading is particularly sensitive to MBS heterogeneity. In terms of the SP/TBA ratio, however, the regression coefficient is remarkably similar whether dealer-customer or all trades are used and whether the dollar trading volume or the number of trades is used in measuring trading activity.

The economic magnitudes are also large. Based on the regression coefficients reported in columns (1)-(3) in Panel A, a one-standard-deviation increase in h^{WAOCS} across coupon cohorts (about 12.58 based on the between standard deviation) is associated with a decrease of about \$62 ($\approx 12.58 \times 4.92$) billion in TBA trading volume and \$4 ($\approx 12.58 \times 0.30$) billion in SP trading volume, and an increase of about 138% ($\approx 12.58 \times 0.11$) in the percentage difference of the SP relative to TBA trading volume.

Table VII. MBS Heterogeneity and Trading Activities

	(1)	(2)	(3)	(4)	(5)	(6)
			A: All Trades			
	Dollar Volume			Number of Trades		
	TBA	SP	SP/TBA	TBA	SP	SP/TBA
h^{WAOCS}	-4.92***	-0.30***	0.11***	-234.71***	-62.15***	0.10***
	(-7.58)	(-2.93)	(7.67)	(-7.10)	(-3.32)	(8.77)
Issuance	10.82***	0.66***	0.02	477.51***	4.58	-0.03***
	(7.01)	(6.22)	(1.57)	(8.47)	(0.21)	(-3.32)
Intercept	290.94***	14.30***	-5.97***	11,519.83***	3,379.04***	-3.88***
	(7.03)	(3.51)	(-16.33)	(6.63)	(3.56)	(-13.06)
Obs	377	377	377	377	377	377
R^2_{adj}	0.85	0.67	0.69	0.86	0.44	0.82
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
			B: Dealer-Customer Trades			
	Dollar Volume			Number of Trades		
	TBA	SP	SP/TBA	TBA	SP	SP/TBA
h^{WAOCS}	-2.16***	-0.25***	0.10***	-43.32***	-44.60***	0.07***
	(-9.11)	(-3.12)	(7.74)	(-6.46)	(-3.54)	(8.41)
Issuance	5.19***	0.64***	0.01*	116.50***	8.17	-0.04***
	(7.66)	(7.66)	(1.67)	(9.20)	(0.58)	(-6.36)
Intercept	105.79***	11.36***	-4.88***	1,761.97***	2,247.09***	-1.65***
	(9.15)	(3.69)	(-16.31)	(6.23)	(3.67)	(-7.88)
Obs	377	377	377	377	377	377
R^2_{adj}	0.85	0.69	0.64	0.84	0.47	0.78
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: In this table we report the results for panel regressions of TBA and SP trading activities as well as their (log) ratios on h^{WAOCS} for FNMA 30-year MBS-S using monthly data. The trading activity is measured both by monthly total par volume (in \$billion) and by monthly total number of trades. The results reported in Panel A include all trades for computing measures of trading activity, while those reported in Panel B include only dealer-customer trades. All regressions control for monthly total new issuance (in \$billion) and time fixed-effects. t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. The overall sample period runs from June 2003 through December 2018 for TBA MBSs and from June 2012 through December 2018 for SP MBSs. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

E. MBS Heterogeneity vs. Prepayment Risk

In this section we conduct analyses to differentiate the effects of MBS heterogeneity from the effects of prepayment risk that have been the main focus of most MBS pricing studies. This is important because our MBS heterogeneity measure is related to prepayment risk, and the OAS measure we use may be related to non-interest-rate prepayment risk premium. Two mechanisms for prepayment risk premium have been proposed in the literature: exposure to market-level prepayment risk and individual-security-level prepayment risk. We address both.

E.1. Premium and Discount Markets

As analyzed in Gabaix et al. (2007) and Diep et al. (2021), exposure to market-level prepayment risk is shown to drive MBS returns, based on a framework in which marginal investors in MBS markets hold specialized aggregate MBS portfolios instead of broadly diversified portfolios. A unique prediction of this framework is that the price of prepayment risk changes signs when the market shifts from one in which premium MBSs dominate (the premium market) to one in which discount MBSs dominate (the discount market). This is because marginal investors holding aggregate MBS market portfolios suffer from an increase in prepayment speed in the premium market, but benefit from it in the discount market.

In contrast, according to our model, MBS heterogeneity always affects MBS yields positively because the effects of MBS heterogeneity arise from the parallel trading environment. Therefore, positive regression coefficients of MBS yields on h_{it}^{WAOCS} in both premium and discount markets would constitute evidence differentiating the effects of MBS heterogeneity from the premium of prepayment risk.

One potential issue with simply running such a regression, however, is that if the relationship between h_{it}^{WAOCS} and prepayment risk exposure changes signs across premium and discount markets, a positive regression coefficient of MBS yields on h_{it}^{WAOCS} in both premium and discount markets may still reflect prepayment risk exposure. To address this issue, in Panel A of Table VIII we report panel regressions of h_{it}^{WAOCS} on moneyness, for the samples of all months, of months when the MBS market is in premium, and of months when the MBS market is in discount, respectively.[32] We find that MBS heterogeneity is always positively depending on moneyness regardless of market type. Given that prepayment risk exposure is monotonic (and decreasing) with mon-

[32]To measure market type, we follow the method of Diep et al. (2021). First, we measure the respective total RPB of all outstanding premium and discount FNMA 30-year MBSs for each month. Then, we classify a month as a discount market when the total RPB for discount securities is greater than the total RPB for premium securities, and as a premium market otherwise. We find that the market has been in premium about 70% of the time during our sample period.

eyness, as shown in Diep et al. (2021), this result implies that the relationship between MBS heterogeneity and prepayment risk exposure is unlikely to change signs across premium and discount markets.

Then we report panel regressions of the OAS on h^{WAOCS} in Panel B of Table VIII, using the samples of all month, of the months when the MBS market is in premium, and of the months when the MBS market is in discount, respectively. We pool all SP groups together again, similar to the study of liquidity shocks in Section IV.C. The regression coefficients on h^{WAOCS} are significantly positive regardless of market type. Compared with regressions with SMM and WAOLTV, the incremental R^2 of h^{WAOCS} is about 4%. Overall, these results show that the effects of MBS heterogeneity are distinct from the effects of exposure to market-level prepayment risk.

E.2. IO and PO Strips

Instead of market-level prepayment factors, many studies focus on individual-security-level prepayment characteristics. Boyarchenko et al. (2019), for example, use IO and PO strips to show that the non-interest-rate prepayment risk premium has significant explanatory power for MBS yields across coupon cohorts. The key feature of IO and PO strips is that their cash flows have opposite exposure to the same prepayment risk (of the same underlying collateral) because prepayments reduce total interest payments while accelerate principal

Table VIII. Premium vs Discount Markets

A: Regression of h^{WAOCS} on Moneyness

	All	Premium Market	Discount Market
Moneyness	9.17***	9.31***	8.45***
	(35.82)	(30.94)	(16.30)
Intercept	16.42***	16.14***	41.10***
	(34.40)	(34.93)	(143.71)
Obs	1,533	1,266	267
R^2_{adj}	0.95	0.95	0.94
Time FE	Yes	Yes	Yes

B: Regression of SP OAS on h^{WAOCS}

	All		Premium Market		Discount Market	
h^{WAOCS}		0.76**		0.73**		0.59*
		(2.51)		(2.31)		(2.22)
SMM	-0.99***	-1.67***	-0.99***	-1.64***	3.86***	0.92
	(-3.62)	(-5.96)	(-3.61)	(-5.73)	(11.79)	(0.66)
WAOLTV	2.62*	0.84	2.76*	0.99	0.86	-0.10
	(1.82)	(0.66)	(1.80)	(0.76)	(1.90)	(-0.10)
Intercept	-148.07	-33.76	-136.91	-21.43	-42.54	32.97
	(-1.56)	(-0.39)	(-1.27)	(-0.23)	(-1.22)	(0.44)
Obs	2,340	2,340	2,280	2,280	60	60
R^2_{adj}	0.47	0.51	0.46	0.51	0.76	0.79
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
SP Type FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: In Panel A we report the results of panel regressions of h^{WAOCS} on moneyness for the samples of all months, months when the MBS market is in premium, and months when the MBS market is in discount, respectively. The market is in premium (discount) in a month when the total RPB of outstanding premium (discount) securities is greater than that of the outstanding discount (premium) securities. In Panel B we report panel regressions of OAS on h^{WAOCS} for all months, premium market months, and discount market months, respectively. We pool all six groups of SP MBSs and include fixed-effects for SP types, while time fixed-effects are included in all regressions as well. OAS regressions include SMM as a control. The t-statistics based on robust standard errors two-way clustered at the time and moneyness cohort dimensions are reported in parentheses (for regressions reported in the last two columns, two-way clustered standard errors cannot be calculated because of the few number of observations, so we only cluster at the moneyness cohort dimension). The overall sample period runs from June 2012 through December 2018. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

payments. We use this feature to differentiate the effects of MBS heterogeneity from that of the individual-security-level prepayment risk: the effects of MBS heterogeneity on returns are positive for both IO and PO strips, while prepayment risk affects returns of IO and PO strips in opposite directions.

In particular, we obtain daily OAS series of IO and PO strips associated with 23 deals or trusts. Their underlying collateral assets are all Fannie Mae 30-year Megas (which are backed by groups of existing pass-through MBSs and/or Megas).[33] For both the IO and PO strips in each trust, we use the average over a month to construct the monthly series. We match them to the sample of MBS characteristics (that are used to measure MBS heterogeneity and reported in Table I) at the cohort-month level. We also obtain characteristics of the collateral securities from eMBS. See Section IA.1 of the Internet Appendix for additional details of the IO/PO strips.

To study how MBS heterogeneity is associated with the OAS of IO/PO strips, we construct monthly OAS series of IO and PO strips at the cohort-month level. Specifically, for each cohort in each month, we take the average of the OAS of the relevant trusts. In Panels A and B

[33] As of June 3, 2019, all TBA-eligible Megas, regardless of issue date, are labeled as "Major Supers". Details are provided at https://capitalmarkets.fanniemae.com/mortgage-backed-securities/structured-transactions-products/supers-and-megas.

of Table IX, we report time-series summary statistics for these monthly OAS series of IO and PO strips for each coupon cohort. The mean OAS of PO strips generally increases from low to high coupon cohorts, ranging from below -60 to above 200 basis points. The mean OAS of IO strips, instead, decreases from 4% to 5% coupon cohorts and then increases from 5% to 7.5% coupon cohorts. The standard deviation of the OAS is larger for IO strips than for PO strips because of their higher price volatility.

Importantly, in Panel C of Table IX, we report panel regressions of the OAS of IO strips (in the first three columns) and of PO strips (in the last three columns) on h_{it}^{WAOCS}. We observe that MBS heterogeneity significantly raises the OAS of both IO and PO strips. The SMM affects the OAS of PO strips significantly but not that of IO strips, and controlling for it does not affect the significance of h_{it}^{WAOCS}. These significant positive effects of MBS heterogeneity on the OAS of both IO and PO strips, which have opposite exposure to the same prepayment risk, constitutes evidence against interpreting our heterogeneity measure as reflecting prepayment risk.

The significant dependence of the OAS of IO/PO strips on MBS heterogeneity is likely because investors can use TBA contracts as a trading option for the underlying collateral MBS of IO/PO strips. This would happen when the value of the underlying collateral MBS is not among the highest; otherwise, it would always be sold on the SP market and

Table IX. OAS of IO/PO Strips and MBS Heterogeneity

A: PO OAS

Coupon	mean	sd	p25	p50	p75
4	-63.48	45.50	-111.20	-54.35	-26.30
4.5	-67.09	30.88	-86.81	-68.46	-46.99
5	6.90	201.90	-71.94	-32.78	17.08
5.5	1.16	196.21	-101.26	-27.66	31.67
6	-4.12	225.10	-99.29	-44.57	15.99
6.5	19.20	219.01	-72.39	-24.31	23.64
7	103.82	295.64	-44.58	9.26	189.45
7.5	244.32	496.08	-40.96	16.40	315.03

B: IO OAS

Coupon	mean	sd	p25	p50	p75
4	544.91	370.08	218.30	371.30	989.57
4.5	413.88	240.00	193.35	359.38	622.71
5	319.43	424.02	6.10	79.98	539.68
5.5	399.84	520.99	4.92	80.58	628.12
6	351.37	524.39	-26.27	39.34	717.59
6.5	351.36	496.66	-10.92	164.51	645.89
7	436.44	592.38	-6.40	246.20	703.00
7.5	553.85	807.05	-37.00	95.68	1076.82

C: Regression of IO and PO OAS

	PO			IO		
h^{WAOCS}	3.98***		8.14***	11.06***		29.77***
	(4.20)		(2.67)	(6.43)		(4.47)
SMM		-27.15***	-25.75***		-30.68	-25.58
		(-3.91)	(-3.89)		(-1.57)	(-1.40)
WAOLTV		20.99***	-12.49		46.73***	-75.73***
		(6.13)	(-1.05)		(5.25)	(-2.81)
Intercept	-207.79***	-1,514.27***	642.81	-472.14***	-3,398.46***	4,492.56***
	(-5.22)	(-6.39)	(0.84)	(-6.16)	(-5.50)	(2.61)
Obs	612	612	612	612	612	612
R^2_{adj}	0.79	0.79	0.79	0.86	0.85	0.87
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

D: Fraction within a Cohort with Higher SMM than IO/PO Collateral

Coupon	mean	sd	min	p25	p50	p75	max
4	0.46	0.10	0.36	0.39	0.41	0.53	0.66
4.5	0.53	0.12	0.40	0.45	0.47	0.67	0.76
5	0.45	0.04	0.38	0.43	0.44	0.48	0.55
5.5	0.47	0.05	0.39	0.43	0.45	0.50	0.65
6	0.40	0.05	0.34	0.37	0.38	0.43	0.62
6.5	0.31	0.07	0.24	0.27	0.28	0.33	0.56
7	0.28	0.05	0.21	0.24	0.27	0.30	0.44
7.5	0.27	0.05	0.19	0.23	0.25	0.30	0.45

Note: Panels A and B report summary statistics for monthly OASs of IO and PO strips of FNMA 30-year MBSs. The average OAS of multiple strips, if available, is used for each cohort in each month. Panel C reports panel regressions of the OAS on h^{WAOCS}, with time fixed-effects included. The t-statistics based on robust standard errors two-way clustered (along the time and coupon dimensions) are reported in parentheses. Panel D reports summary statistics for the monthly time-series of the fraction of outstanding MBSs that have higher SMM than that of the IO/PO collateral for each cohort. The overall sample period runs from January 2004 through April 2012. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

its price would not depend on MBS heterogeneity (see Proposition 2). To provide some supportive evidence, for each cohort in each month, we take the average of the SMM of all available IO/PO collateral MBSs. We then match these collateral MBSs to the whole sample of CUSIP-level MBS characteristics and compute, for each cohort in each month, the fraction of outstanding MBSs with higher SMM than the IO/PO collateral. Panel D of Table IX report time-series summary statistics of this fraction for each coupon cohort. The median fractions are all below 50%, and even lower than 30% for 6.5%-7.5% cohorts. That is, the IO/PO collateral fall within the lower range of the value distribution within a cohort indeed. Thus, they are likely to be delivered into TBA contracts when SP trading cost is high to sellers.

V. Conclusion

To the best of our knowledge, this paper conducts the first analysis of the distinctive asset pricing effects of the TBA/SP parallel trading environment. In particular, we construct a simple "liquidity-based asset pricing" model that allows investors to choose between TBA and SP trading. Measuring the dispersion of MBS values within a cohort based on individual-MBS-level prepayment characteristics, denoted as MBS heterogeneity, we empirically confirm the important effects of MBS heterogeneity on pricing and trading implied from the model. We also provide evidence to distinguish the effects of MBS heterogeneity from the impacts of prepayment risks.

The agency MBS market is of broad interest given its important role in the U.S. financial system, and so are the cohort-based TBA trading mechanism and the economic effects of MBS heterogeneity. A TBA-like trading mechanism can be potentially applied to most OTC fixed-income markets (Spatt (2004), Bessembinder et al. (2019), and Gao et al. (2017)). Further understanding of these market design issues can be achieved built on the economic effects we document here.

Appendix A. Proofs

Proof of Proposition 1. At time 2, the seller of an asset with value v_k may obtain revenue of $P_2^{\mathrm{SP}}(v_k) - C_2^{\mathrm{SP}} = v_k - C_2^{\mathrm{SP}}$ in the SP market or $P_2^{\mathrm{TBA}} = v_m - h_d$ in the TBA market. Hence, she chooses the TBA market if and only if $v_m - h_d \geq v_k - C_2^{\mathrm{SP}}$, which is equivalent to $v_k \leq v_m - h_d + C_2^{\mathrm{SP}} = \bar{v}_2$. □

Proof of Lemma 1. At time 2, C_2^{SP} is realized and an MBS of value v_k

can generate revenue of

$$\max\left\{v_k - C_2^{\mathrm{SP}}, v_m - h_d\right\} = \begin{cases} v_k - C_2^{\mathrm{SP}} & \text{if } v_k > \bar{v}_{2,h} \\ v_k - c_{2,\ell} & \text{if } v_k \in [\bar{v}_{2,\ell}, \bar{v}_{2,h}] \text{ and } C_2^{\mathrm{SP}} = c_{2,\ell} \\ v_m - h_d & \text{if } v_k \in [\bar{v}_{2,\ell}, \bar{v}_{2,h}] \text{ and } C_2^{\mathrm{SP}} = c_{2,h} \\ v_m - h_d & \text{if } v_k < \bar{v}_{2,\ell} \end{cases} \tag{A1}$$

Hence, at time 1, the buyer is willing to pay

$$\begin{aligned} P_1^{\mathrm{SP}}(v_k) &= (1-\rho)v_k + \rho\,\mathrm{E}\left[\max\left\{v_k - C_2^{\mathrm{SP}}, v_m - h_d\right\}\right] \\ &= v_k - \rho \times \begin{cases} \mathrm{E}[C_2^{\mathrm{SP}}] & \text{if } v_k > \bar{v}_{2,h}, \\ \pi_h(v_k - v_m + h_d) + (1-\pi_h)c_{2,\ell} & \text{if } v_k \in \left[\bar{v}_{2,\ell}, \bar{v}_{2,h}\right], \\ v_k - v_m + h_d & \text{if } v_k < \bar{v}_{2,\ell}. \end{cases} \end{aligned} \tag{A2}$$

□

Proof of Proposition 2. At time 1, because an MBS with value v_k generates $P_1^{\mathrm{SP}}(v_k) - C_1^{\mathrm{SP}}$ in the SP market and $P_1^{\mathrm{TBA}} = v_m - h_d$ in the TBA market, a seller is indifferent between the TBA and the SP market if the value of her MBS $\bar{v}_1$ satisfies $P_1^{\mathrm{TBA}} = P_1^{\mathrm{SP}}(\bar{v}) - C_1^{\mathrm{SP}} = \bar{v}_1 - (\bar{v}_1 - P_1^{\mathrm{SP}}(\bar{v}_1)) - C_1^{\mathrm{SP}}$, which implies that $\bar{v}_1 = P_1^{\mathrm{TBA}} + C_1^{\mathrm{SP}} + (\bar{v}_1 - P_1^{\mathrm{SP}}(\bar{v}_1))$. Lemma 1 then implies that

$$\bar{v}_1 = v_m - h_d + C_1^{\mathrm{SP}} + \rho \times \begin{cases} \mathrm{E}[C_2^{\mathrm{SP}}] & \text{if } \bar{v}_1 > \bar{v}_{2,h}, \\ \pi_h(\bar{v}_1 - v_m + h_d) + (1-\pi_h)c_{2,\ell} & \text{if } \bar{v}_{2,\ell} \le \bar{v}_1 \le \bar{v}_{2,h}, \\ \bar{v}_1 - v_m + h_d & \text{if } \bar{v}_1 < \bar{v}_{2,\ell}. \end{cases} \tag{A3}$$

It follows that

$$\bar{v}_1 = v_m - h_d + \begin{cases} C_1^{\mathrm{SP}} + \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] & \text{if } C_1^{\mathrm{SP}} + \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] > c_{2,h}, \\ \frac{C_1^{\mathrm{SP}} + \rho(1-\pi_h)c_{2,\ell}}{1-\rho\pi_h} & \text{if } c_{2,\ell} \le \frac{C_1^{\mathrm{SP}} + \rho(1-\pi_h)c_{2,\ell}}{1-\rho\pi_h} \le c_{2,h}, \\ \frac{C_1^{\mathrm{SP}}}{1-\rho} & \text{if } \frac{C_1^{\mathrm{SP}}}{1-\rho} < c_{2,\ell}, \end{cases} \tag{A4}$$

which can be rewritten as (8). Because $P_1^{\mathrm{SP}}(v_k)$ increases with v_k, then an MBS with value less than $\bar{v}_1$ should be sold in the TBA market and the SP market otherwise. □

Proof of Corollary 1. If $C_1^{\mathrm{SP}} > c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$, then Proposition 2 implies that $\bar{v}_1 = v_m - h_d + C_1^{\mathrm{SP}} + \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] > v_m - h_d + c_{2,h} = \bar{v}_{2,h}$. Because the value v_k of any time-1 SP MBS is greater than $\bar{v}_1$, Lemma 1 implies that $P_1^{\mathrm{SP}}(v_k) = v_k - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$, which is independent from h_d.

If $C_1^{\mathrm{SP}} < (1-\rho)c_{2,\ell}$, then Proposition 2 implies that $\bar{v}_1 = v_m - v_d + \frac{C_1^{\mathrm{SP}}}{1-\rho} <$

$v_m - v_d + c_{2,\ell} = \bar{v}_{2,\ell} \le \bar{v}_{2,h}$. If $(1-\rho)c_{2,\ell} \le C_1^{\mathrm{SP}} \le c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$, then

$$\frac{C_1^{\mathrm{SP}} + \rho(1-\pi_h)c_{2,\ell}}{1-\rho\pi_h} \le \frac{c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}] + \rho(1-\pi_h)c_{2,\ell}}{1-\rho\pi_h} = c_{2,h}. \tag{A5}$$

Hence $\bar{v}_2 \in [v_{2,\ell}, v_{2,h}]$. Thus, when $C_1^{\mathrm{SP}} \le c_{2,h} - \rho\,\mathrm{E}[C_2^{\mathrm{SP}}]$, $\bar{v}_1 \le \bar{v}_{2,h}$ and Lemma 1 implies that $P_1^{\mathrm{SP}}(v_k)$ decreases with h_d. $\square$

REFERENCES

Akerlof, George A., 1970, The market for lemons: Quality uncertainty and the market mechanism, *Quarterly Journal of Economics* 84, 488–500.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223 – 249.

Amihud, Yakov, Haim Mendelson, and Lasse Heje Pedersen, 2006, Liquidity and asset prices, *Foundations and Trends in Finance* 1, 269–364.

An, Yu, Wei Li, and Zhaogang Song, 2020, Cohort trading and security design: Theory and evidence from agency mbs markets, *Johns Hopkins Carey Business School, working paper* .

Bartolini, Leonardo, Spence Hilton, Suresh Sundaresan, and Christopher Tonetti, 2011, Collateral values by asset class: Evidence from primary securities dealers, *Review of Financial Studies* 24(1), 248–278.

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

Bessembinder, Hendrik, Chester Spatt, and Kumar Venkataraman, 2019, A survey of the microstructure of fixed-income markets, *Journal of Financial and Quantitative Analysis* .

Boudoukh, Jacob, Matthew Richardson, Richard Stanton, and Robert Whitelaw, 1997, Pricing mortgage-backed securities in a multifactor interest rate environment: A multivariate density estimation approach, *Review of Financial Studies* 10, 405–446.

Boyarchenko, Nina, Andreas Fuster, and David O. Lucca, 2019, Understanding mortgage spreads, *The Review of Financial Studies* 32, 3799–3850.

Brennan, Michael J., and Eduardo S. Schwartz, 1985, Determinants of gnma mortgage prices, *Real Estate Economics* 13, 209–228.

Chen, Jiakai, Haoyang Liu, Asani Sarkar, and Zhaogang Song, 2021, Dealers and the dealer of last resort: Evidence from mbs markets in the covid-19 crisis, *Federal Reserve Bank of New York Staff Reports* 933.

Chen, Si, 1996, Understanding option-adjusted spreads, *The Journal of Portfolio Management* 22, 104–113.

Chernov, Mikhail, Brett R. Dunn, and Francis A. Longstaff, 2017, Macroeconomic-Driven Prepayment Risk and the Valuation of Mortgage-Backed Securities, *The Review of Financial Studies* 31, 1132–1183.

Cheyette, Oren, 1996, Implied prepayments, *The Journal of Portfolio Management* 23, 107–115.

Deng, Yongheng, John M. Quigley, and Robert Van Order, 2000, Mortgage terminations, heterogeneity and the exercise of mortgage options, *Econometrica* 68, 275–307.

Diep, Peter, Andrea L. Eisfeldt, and Scott Richardson, 2021, The cross section of mbs returns, *The Journal of Finance* forthcoming.

Downing, Christopher, Dwight Jaffee, and Nancy Wallace, 2009, Is the market for mortgage backed securities a market for lemons?, *Review of Financial Studies* 22–7, 2457–2494.

Downing, Christopher, Richard Stanton, and Nancy Wallace, 2005, An empirical test of a two-factor mortgage valuation model: How much do house prices matter?, *Real Estate Economics* 33, 681–710.

Duarte, Jefferson, 2007, The Causal Effect of Mortgage Refinancing on Interest Rate Volatility: Empirical Evidence and Theoretical Implications, *The Review of Financial Studies* 21, 1689–1731.

Duarte, Jefferspm, Francis Longstaff, and Fan Yu, 2007, Risk and return in fixed income arbitrage: Nickels in front of a steamroller?, *Review of Financial Studies* 20, 769–811.

Dunn, Kenneth, and John McConnell, 1981, Valuation of gnma mortgage-backed securities, *Journal of Finance* 36, 599–616.

Dunn, Kenneth B., and Chester S. Spatt, 2005, The effect of refinancing costs and market imperfections on the optimal call strategy and the pricing of debt contracts, *Real Estate Economics* 33, 595–617.

Easley, David, and Maureen O'Hara, 2003, Chapter 17 microstructure and asset pricing, in *Financial Markets and Asset Pricing*, volume 1 of *Handbook of the Economics of Finance*, 1021–1051 (Elsevier).

Fabozzi, Frank, and Steven Mann, 2011, Handbook of fixed income securities, *McGraw-Hill; 8 edition* .

FRBNY, 2020, Statement regarding treasury securities, agency mortgage-backed securities, and repurchase agreement operations.

Friewald, Nils, Rainer Jankowitsch, and Marti G. Subrahmanyam, 2017, Transparency and Liquidity in the Structured Product Market, *The Review of Asset Pricing Studies* 7, 316–348.

Frost, Josh, Lorie Logan, Antoine Martin, Patrick McCabe, Fabio Natalucci, and Julie Remache, 2015, Overnight rrp operations as a monetary policy tool: Some design considerations, *Finance and Economics Discussion Series* 2015-010.

Gabaix, Xavier, Arvind Krishnamurthy, and Olivier Vigneron, 2007, Limits of arbitrage: Theory and evidence from the mortgage-backed securities market, *Journal of Finance* 62, 557–595.

Gao, Pengjie, Paul Schultz, and Zhaogang Song, 2017, Liquidity in a market for unique assets: Specified pool and tba trading in the mortgage backed securities market, *Journal of Finance* 72-3, 1119–1170.

Gao, Pengjie, Paul Schultz, and Zhaogang Song, 2018, Trading methods and trading costs for agency mortgage backed securities, *Journal of Investment Management* 16, 29–46.

Gay, Gerald, and Steven Manaster, 1984, The quality delivery implicit in futures contracts, *Journal of Financial Economics* 13, 353–370.

Gurkaynak, Refet, Brian Sack, and Jonathan Wright, 2007, The u.s. treasury yield curve: 1961 to the present, *Journal of Monetary Economics* 54(8), 2291–2304.

Hansen, Samuel, 2014, Mortgage convexity, *Journal of Financial Economics* 113-2, 270–299.

Hayre, Lakhbir, 2001, Salomon smith barney guide to mortgage-backed and asset-backed securities, *Wiley* .

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

He, Zhiguo, Paymon Khorrami, and Zhaogang Song, 2019, Commonality in credit spread changes: Dealer inventory and intermediary distress, Working Paper 26494, National Bureau of Economic Research.

He, Zhiguo, and Zhaogang Song, 2021, Agency mbs as safe assets, *working paper* Johns Hopkins Carey Business School.

Hegde, S, 1988, An empirical analysis of implicit delivery options in the treasury bond futures contract, *Journal of Banking and Finance* 12, 469–492.

Hemler, Michael, 1990, The quality delivery option in treasury bond futures contracts, *Journal of Finance* 45-5, 1565–1586.

Himmelberg, Charles, Marty Young, Hui Shan, and Chris Henson, 2013, The impact of fed mbs purchases on tba pool quality, *Goldman Sachs* .

Hu, Xing, Jun Pan, and Jiang Wang, 2013, Noise as information for illiquidity, *Journal of Finance* 68, 2223–2772.

Huh, Yesol, and You Suk Kim, 2020, Cheapest-to-deliver pricing and endogenous mbs heterogeneity, *Working Paper* Board of Governors of the Federal Reserve System.

Kane, A., and A. Marcus, 1986, The quality option in the treasury bond futures market: An empirical assessment, *Journal of Futures Marketse* 6, 231–248.

Kim, You Suk, and Yesol Huh, 2019, The real effects of the secondary market trading structure: Evidence from the mortgage market, *working paper* .

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, 2013, The ins and outs of lsaps, *Working paper* .

Levin, Alexander, and Andrew Davidson, 2005, Prepayment risk-and option-adjusted valuation of mbs, *The Journal of Portfolio Management* 31, 73–85.

Li, Wei, and Zhaogang Song, 2020, Asset heterogeneity, market fragmentation, and quasi-consolidated trading, *working paper* .

Liu, Haoyang, Zhaogang Song, and James Vickery, 2021, Defragmenting markets: Evidence from agency mbs, *Federal Reserve Bank of New York Staff Reports, no. 965* .

Longstaff, Francis A., 2005, Borrower credit and the valuation of mortgage-backed securities, *Real Estate Economics* 33, 619–661.

Malkhozov, Aytek, Philippe Mueller, Andrea Vedolin, and Gyuri Venter, 2016, Mortgage Risk and the Yield Curve, *The Review of Financial Studies* 29, 1220–1253.

Richard, Scott F., and Richard Roll, 1989, Prepayments on fixed-rate mortgage-backed securities, *The Journal of Portfolio Management* 15, 73–82.

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

Schwartz, Eduardo, and Walter Torous, 1989, Prepayment and the valuation of mortgage-backed securities, *Journal of Finance* 44, 375–392.

Song, Zhaogang, and Haoxiang Zhu, 2018, Quantitative easing auctions of treasury bonds, *Journal of Financial Economics* 128, 103 – 124.

Song, Zhaogang, and Haoxiang Zhu, 2019, Mortgage Dollar Roll, *The Review of Financial Studies* 32, 2955–2996.

Spatt, Chester, 2004, Frictions in the bond market, *Keynote Speech: Second MTS Conference on Financial Markets* .

Stanton, Richard, 1995, Rational prepayment and the valuation of mortgage-backed securities, *Review of Financial Studies* 8, 677–708.

Stanton, Richard, and Nancy Wallace, 1998, Mortgage choice: What's the point?, *Real Estate Economics* 26, 173–205.

Svensson, Lars, 1994, Estimating and interpreting forward rates: Sweden 1992-1994, *National Bureau of Economic Research Working Paper #4871* .

Titman, Sheridan, and Walter Torous, 1989, Valuing commercial mortgages: An empirical investigation of the contingent-claims approach to pricing risky debt, *Journal of Finance* 44, 345–373.

Vayanos, Dimitri, and Jiang Wang, 2013, Chapter 19 market liquidity: Theory and empirical evidence, *Handbook of the Economics of Finance* 2, 1289 – 1361.

Vickery, James, and Joshua Wright, 2013, Tba trading and liquidity in the agency mbs market, *Federal Reserve Bank of New York Economic Policy Review* 19.

Internet Appendix for
"Asset Pricing with Cohort-Based Trading in MBS Markets"

This Internet Appendix provides additional results and robustness checks for the paper "Asset Pricing with Cohort-Based Trading in MBS Markets," by Nicola Fusari, Wei Li, Haoyang Liu, and Zhaogang Song.[1]

[1]Citation format: Nicola Fusari, Wei Li, Haoyang Liu, and Zhaogang Song, Internet Appendix for "Asset Pricing with Cohort-Based Trading in MBS Markets," *Journal of Finance*.

IA.1. Additional Details of the Data and Measures

In this section, we provide additional details of the data and measures that are mentioned briefly in the paper.

First, the MBS characteristics such as WAOCS, WAOLTV, RPB, and so on, are calculated based on the "Fixed-Rate Quartile" disclosure files that Fannie Mae began to release in June 2003. These MBS-level characteristics are calculated using values for individual loans at the time of MBS issuance weighted by the remaining loan balance at the time of calculation. For example, the FICO scores used to compute WAOCS underlying an MBS are credit scores at issuance rather than credit scores at the time of calculation. But the loan balance used as the weight is the remaining loan balance as of the release date of the disclosure files for each month, so there is time-series variation in WAOCS for an MBS. The disclosure files are released mostly on the fourth business day of each month.[2] We follow the procedure as described in Himmelberg et al. (2013) and also used in Song and Zhu (2019) to exclude the set of MBSs that are least likely to be delivered into TBA contracts as follows: For each coupon cohort in each month, we eliminate MBSs that have at least one of the following characteristics: a refinance share greater than 75%, a RPB less than $150,000, a WAOLTV above 85%, and a WAOCS

[2] Details on the disclosure are available at https://www.fanniemae.com/media/16486/display.

below 680. MBSs with these characteristics "that inhibit efficient prepayments command a price premium, and are not delivered into TBAs" (Himmelberg et al. (2013)).

Second, our baseline measure of MBS heterogeneity uses WAOCS as the prepayment characteristic. To check the effect of WAOCS, as well as that of other prepayment characteristics, on prepayment rates, Table IA.I reports results of regressing prepayment rates on characteristics for newly issued Fannie Mae 30-year TBA-eligible MBS from January 2011 through December 2018. The dependent variables are average prepayment rates in the first 6 months (SMM^{6m}), 12 months (SMM^{12m}), 24 months (SMM^{24m}), and 36 months (SMM^{12m}) since issuance, while the independent variables are WAOCS, WAOLTV, and WAOSIZE at issuance.

Each of these three characteristics is a key input for prepayment models, with the appealing feature that their relationship with prepayment rate of in-the-money MBSs is largely monotonic (Fabozzi and Mann (2011)).[3] The effect of WAOCS on prepayment rate is positive because high FICO borrowers can better exploit refinance opportunities. The effect of WAOSIZE is also positive because savings from refinancing larger loans are higher and more likely to outweigh certain fixed costs of refinancing. The effect of WAOLTV on prepayment risk is typically

[3]During this sample period, most of newly-issued MBSs are in the money, so MBSs with higher prepayment rates are those with higher prepayment risks.

negative because higher-LTV loans are more likely to be underwater and hence less likely to refinance. Higher-LTV loans, however, are also more likely to default and end up with full prepayment by the guaranteeing agency. In general, the effect of WAOLTV on prepayment rate is negative in the short horizon after issuance when default is not of much relevance, but can turn positive in the long horizon when default becomes more likely (Fabozzi and Mann (2011)).

The first three columns of Table IA.I reports the results of regressing $\text{SMM}^{6\text{m}}$ on WAOCS, WAOLTV, and WAOSIZE, respectively. We include coupon cohort fixed effects, so the regression coefficients capture how these prepayment characteristics affect prepayment rates fixing a coupon cohort. We observe that all these characteristics significantly affect prepayment rates with expected signs. In terms of explanatory power, a regression with only coupon fixed effects produces an adjusted R^2 of 2.2%, relative to which WAOCS, WAOSIZE, and WAOLTV bring in an incremental explanatory power of 0.9%, 3.6%, and 0.1% (note that these are individual-MBS-level regressions, so R^2s are unsurprisingly low with a single regressor). The third column reports multivariate regressions with all the three characteristics, which achieve a R^2 of about 6% for $\text{SMM}^{6\text{m}}$. The explanatory power increases further to 17%, 27%, and 28.5% for $\text{SMM}^{12\text{m}}$, $\text{SMM}^{24\text{m}}$, and $\text{SMM}^{36\text{m}}$ respectively. We also note that the signs of WAOCS and WAOSIZE are always significantly positive across different horizons, but that of WAOLTV turn positive in longer horizons, as expected.

Overall, the WAOCS measure we use in the baseline analysis does significantly affect prepayment rates. Although we do not use WAOSIZE that has the highest explanatory power for prepayment rate as the main prepayment characteristic, to alleviate concerns on cherry-picking, we conduct robustness checks using heterogeneity measures based on WAOSIZE, WAOLTV, and so on below (see Table IA.VII and Table IA.VIII).

Third, in Panel A of Table IA.II we report the time-series summary statistics of h_{it}^{WAOCS} for each coupon cohort i in our sample. The mean dispersion in WAOCS increases monotonically from low to high coupon cohorts, ranging from approximately 16 to 48. This pattern arises because, as Panel C of Table I shows, the mean of 50th percentiles and the mean of 5th percentiles both decrease from low to high coupon cohorts but the former decreases faster than the latter. In fact, the mean of 50th percentiles drops by 67 ($\approx 775 - 708$) while the mean of the 95th percentiles drops only by 36 ($\approx 792 - 756$). Again, this is, as discussed in Section III, consistent with the fact that high-FICO loans are refinanced more quickly into low-coupon MBSs in the context of falling mortgage rates during our sample period. Moreover, the time series variation of h_{it}^{WAOCS} seems to be low, especially for low-coupon cohorts. In Panel B of Table IA.II we report the time-series summary statistics for h_{it}^{SMM} for each coupon cohort i included in our sample. By construction, the dispersion measure has a theoretical upper bound of 100% because prepayment rates are bounded between 0% and 100%. We observe that the average value of h^{SMM} increases from low to high coupon cohorts.

Fourth, in Table IA.III, we provide summary statistics for IO/PO trusts that are used in Section IV.E of the paper. The overall sample covers coupon cohorts of 4%-7.5% from January 2004 through April 2012. The 4%, 7%, and 7.5% cohorts each contain a single trust, while other cohorts contain three to five trusts. The trusts are large, mostly with notional value greater than $2 billion. The vintage is between 2000 and 2010, except one trust issued in 1994 and another issued in 1999. The FICO scores are lower for trusts with higher coupon rates. This pattern is also documented in our whole sample MBS characteristics as reported in Table I. The WAC is usually higher than the cohort coupon rate by about 50 basis points, while the LTV ratio ranges between 68% and 80%. The time series mean of moneyness is between -0.16 and 2.11 and that of SMM is between 1.23% and 2.98%.

IA.2. Empirical Relevance of Equilibrium Types

In this section, we calibrate the model and show empirical evidence that SP pricing does depend on future TBA trading cost on most days.

We relax the assumption of the baseline model that the SP trading costs follow a simple two-point distribution. In particular, let C_t^{SP} represent the SP selling cost at time t. We assume that $C_t^{\text{SP}} \overset{\text{iid}}{\sim} F_c$ so that

$\bar{c} = \mathrm{E}[C_t^{\mathrm{SP}}]$ and $C_t^{\mathrm{SP}} \in [c_{\min}, c_{\max}]$. Then, generalizing Corollary 1, prices of some SP MBS at time t depend on h_d if

$$C_t^{\mathrm{SP}} \leq c_{\max} - \rho_\delta \, \mathbf{E}[C_{t+\delta}^{\mathrm{SP}}] = c_{\max} - \rho_\delta \bar{c}, \tag{IA.1}$$

where $\rho_\delta \leq 1$ equals the probability of reselling at time $t+\delta$. Under this condition, some MBSs sold on the SP market at time t may be resold in the TBA market at time $t+\delta$. If C_t^{SP} is drawn repeatedly from F_c, the condition (IA.1) holds with probability

$$\Pr\left\{C_t^{\mathrm{SP}} \leq c_{\max} - \rho_\delta \bar{c}\right\} = F_c\left(c_{\max} - \rho_\delta \bar{c}\right). \tag{IA.2}$$

We then empirically estimate the *lower bound* of (IA.2) by setting $\rho_\delta = 1$. We first compute the SP trading cost for a coupon cohort on each day as follows. For each MBS j within a coupon cohort i traded in the SP market on day t, we split all trades into three types: inter-dealer trades, dealer purchases from customer, and dealer sales to customer. We compute the volume-weighed average prices for these three types, denoted as P_{ijt}^{ID}, P_{ijt}^{DSale}, and P_{ijt}^{DBuy}. We then calculate the SP trading cost as $\log(P_{ijt}^{\mathrm{DSale}}/P_{ijt}^{\mathrm{ID}})$ when only P_{ijt}^{DSale} is present, $-\log(P_{ijt}^{\mathrm{DBuy}}/P_{ijt}^{\mathrm{ID}})$ when only P_{ijt}^{DBuy} is present, and the average of the two log differences when both are present (see Hendershott and Madhavan (2015) for similar measures). Then, we compute the average of the trading costs of all SP MBSs weighted by the total trading volume on day t for each coupon cohort i on each day t.

We report in the first four columns of Table IA.IV the number of days, mean, 95th percentile, and 90th percentile of the daily time series of SP trading costs for each coupon cohort from May 16, 2011 to August 5, 2015. We observe that the SP trading cost is about 14 bps for the 3% and 3.5% coupon cohorts, which are in active issuance in this sample period. The trading cost increases to about 30 bps for the 5% and 5.5% coupon cohorts and to over 40 bps for the 6% and 6.5% coupon cohorts that are deeply seasoned.

Most importantly, the last two columns of Table IA.IV report the fraction of days when the SP trading cost C_t^{SP} is below $C_{95\%}^{\mathrm{SP}} - \bar{c}$ or $C_{90\%}^{\mathrm{SP}} - \bar{c}$, where $C_{95\%}^{\mathrm{SP}}$, $C_{90\%}^{\mathrm{SP}}$, and $\bar{c}$ equal the 95th percentile, 90th percentile, and the mean of the SP cost, respectively. That is, we empirically calculate $\Pr\{C_t^{\mathrm{SP}} \le C_{95\%}^{\mathrm{SP}} - \bar{c}\}$ and $\Pr\{C_t^{\mathrm{SP}} \le C_{90\%}^{\mathrm{SP}} - \bar{c}\}$. Because $c_{\max} \ge C_{95\%}^{\mathrm{SP}} \ge C_{90\%}^{\mathrm{SP}}$ and the probability of reselling $\rho_\delta \le 1$, we have that $c_{\max} - \rho_\delta \bar{c} \ge C_{95\%}^{\mathrm{SP}} - \bar{c} \ge C_{90\%}^{\mathrm{SP}} - \bar{c}$. Thus, the empirical measures $\Pr\{C_t^{\mathrm{SP}} \le C_{95\%}^{\mathrm{SP}} - \bar{c}\}$ and $\Pr\{C_t^{\mathrm{SP}} \le C_{90\%}^{\mathrm{SP}} - \bar{c}\}$ are conservative estimates of the true likelihood of SP pricing being affected by future TBA trading. We find that the estimated likelihoods are fairly large: $\Pr\{C_t^{\mathrm{SP}} \le C_{95\%}^{\mathrm{SP}} - \bar{c}\}$ exceeds 80% for all coupon cohorts and $\Pr\{C_t^{\mathrm{SP}} \le C_{90\%}^{\mathrm{SP}} - \bar{c}\}$ ranges between 57% and 77%. Moreover, the likelihoods are larger for higher coupon cohorts, which tend to be more heterogeneous. Given the conservativeness of our estimation method, these results suggest that in practice SP pricing depends on future TBA trading costs on *at least* 80%, and probably more, of trading days.

IA.3. Burnout Effects

In this section, we study the effects of burnout, which captures the path dependence of mortgage prepayment. Following Schwartz and Torous (1993), for each MBS j within a coupon cohort i at month t, we compute a burnout measure as

$$\text{Burnout}_{ijt} = \sum_{\tau=\tau^{\text{Issuance}}}^{t} \max\{\log(c_i/c_t), 0\}, \tag{IA.3}$$

where c_i is the coupon rate of this MBS and c_t is the current coupon rate. That is, the burnout measure captures accumulates the moneyness of the prepayment option over time. As our analysis is at the coupon-cohort level, we then take the median of Burnout_{ijt} among all MBSs within the cohort i at month t as the burnout measure at the cohort level, denoted as Burnout_{it}.

As mentioned in Section III, the lower SMM of the 7%-cohort when compared with the slightly lower coupon cohorts as reported in Table I is consistent with a burnout effect. Indeed, we find that the time series mean of Burnout_{it} for the 7% cohort is about 40, higher than that for the 6.5% cohort about 33. We further conduct a more direct test of the effects of burnout on prepayment rates. In particular, we regress

SMM_{it}, which is measured as the median of prepayment rates of different MBSs within cohort i at month t, on Burnout_{it}, controlling for the coupon cohort FE. From the first column in Panel A of Table IA.V, we observe that burnout negatively affects prepayment rates significantly, indeed. Further, the second column reports the result of the regression including the interaction term between Burnout_{it} and coupon rate. We observe that its regression coefficient is significantly negative, implying that the effect of burnout is stronger for cohorts of higher coupons.

We finally construct a heterogeneity measure using burnout, h^{Burnout}. Panel B of Table IA.V reports the results when including h^{Burnout} as a control to (14). We observe that h^{Burnout} has explanatory power for OAS of some SP groups. Importantly, the effects of h^{WAOCS} are robust when including h^{Burnout} as a control.

IA.4. Robustness Checks

In this section, we report the results of a number of robustness checks.

First, the data sample in the main analysis of Section IV excludes cohorts with fewer than 1,000 MBSs. In Panel A of Table IA.VI, we report all the main results (those reported in Table IV, Table V, Table VI, and Table VII) for the sample excluding coupon cohorts with fewer than 2,000 MBSs. We observe that regression coefficients on MBS heterogeneity are highly significant, like those in the main analysis.

Second, we report in Panel B of Table IA.VI the main results using the sample of Freddie Mac, rather than Fannie Mae, 30-year MBSs. The results are similar to those obtained using Fannie Mae 30-year MBSs in the main analyses.

Third, we construct alternative measures of MBS heterogeneity. In particular, Panels A and B of Table IA.VII report results of our main analysis using $h^{\text{WAOCS},10\%}$ (the difference between the 90th percentile and the median of WAOCS) and h^{WAOSIZE} (the difference between the 95th percentile and the median of WAOSIZE), respectively. Moreover, Panels A and B of Table IA.VIII report results of our main analysis using h_{it}^{SMM} defined in (12) and h_{it}^{Combine} (defined as the fitted value in regressions of $h_{i,t+1}^{\text{SMM}}$ on h_{it}^{WAOCS}, h_{it}^{WAOSIZE}, and h_{it}^{WAOLTV}), respectively. Our main results are robust using all these alternative measures of MBS heterogeneity. In fact, we find that h_{it}^{Combine} works best, as expected, but we do not use it as the baseline measure given its look-ahead bias. Instead, we use h_{it}^{WAOCS} because it is free from look-ahead bias, performs well, and, importantly, is simple to interpret.

Fourth, a few studies, such as Fabozzi and Mann (2011) and Belikoff et al. (2010), argue that the OAS based on the LIBOR swap curve may be a better measure in practice because LIBOR is widely used as the benchmark borrowing rate and swap rates are quoted more uniformly and densely than Treasury yields. In Panel A of Table IA.IX we repeat the main analyses (those reported in Table V and Table VI) using an

OAS series based on the LIBOR-swap curve. The results remain nearly unchanged.

Fifth, as discussed above, the OAS series used in our main analysis depends on a dealer's prepayment model that may be mis-specified. To alleviate this concern, we obtain cohort-level monthly OAS series of SP MBSs from another major Wall Street MBS dealer (OAS series of various SP MBSs groups within a cohort are not available from this dealer). We also obtain a series of *hedged returns* for SP MBSs from this dealer, which are favored by some studies such as Diep et al. (2021). Regression results for these alternative OAS and return series are reported in Panel B of Table IA.IX. The robust effects of MBS heterogeneity on SP yields reported mitigate concerns regarding the impact of prepayment model mis-specifications on our main findings.

Sixth, Table IA.X reports regression results of MBS trading activities on h^{WAOCS}, by controlling for outstanding balance instead of new issuance. The results are similar.

Seventh, we conduct robustness checks on dispersion in dealer prepayment forecasts. We obtain prepayment rate forecasts that major Wall Street dealers provide to Bloomberg monthly. The prepayment speeds are quoted for generic coupon cohorts according to the PSA convention in which the annualized CPR is adjusted for the age of the underlying mortgages. The forecasts are given for several interest rate scenarios, ranging from 300 basis points below the current rate to 300 basis points above. Following Carlin et al. (2014), for each coupon cohort i in each month t, we take the ratio of each dealer d's PSA prepayment forecast for the -100 basis point scenario to that for the $+100$ basis point scenario. That is, this ratio measures the relative change in prepayment forecast as interest rates move from 100 basis points above the current level of interest rates to 100 basis points below, which captures the sensitivity of prepayments to changes in interest rates. We then compute the simple standard deviation of the ratios across dealers who provide forecasts for coupon i in month t, denoted as *Dealer Dispersion*. This dispersion measure is calculated for the coupon cohort with at least three dealers' forecasts available.

As shown in the last column of Panel A of Table IA.XI, we find that *Dealer Dispersion* decreases from low-coupon cohorts to high-coupon cohorts. This decreasing pattern of *Dealer Dispersion* across coupon cohorts is consistent with lower prepayment uncertainty of deeper in-the-money cohorts: dealers are more sure about their high prepayment rates and hence disagree less.[4] That is, *Dealer Dispersion* captures the

[4]If we only consider interest-rate-driven refinancing, it is expected that dealer dispersion should be larger for at-the-money cohorts than for both in-the-money and out-of-the-money cohorts. But because out-of-the-money cohorts have low outstanding balance and very inactive trading, our data sample mainly consists of at-the-money and in-the-money cohorts for which a monotonically decreasing pattern is found in regressions of dealer dispersion on coupon rate. Furthermore, consider prepayments driven by employment-related relocation, cash-out refinancing, and so on, which are

dispersion in dealers' forecasts of the *cohort-level* average prepayment rate, which differs distinctively from MBS heterogeneity we focus on that captures the prepayment dispersion of *individual* MBSs within a coupon cohort. One the one hand, even when all dealers agree on the cohort-level average prepayment rate, individual MBSs can still exhibit large heterogeneity in prepayment rates; on the other hand, even when MBSs are homogeneous within each cohort, dealers may still differ in forecasts of average prepayment rate because of their differences in interest rate forecasts and prepayment models (Carlin et al. (2014)).

The decreasing pattern of *Dealer Dispersion* across coupon cohorts leads to a statistical negative relationship between our MBS heterogeneity measure h^{WAOCS} and *Dealer Dispersion* (recall that h^{WAOCS} increases across coupon cohorts). However, we find that *Dealer Dispersion* does not seem to affect MBS yields as economic theories predict. In particular, as a measure of prepayment uncertainty, *Dealer Dispersion* is expected to *positively* affect MBS yields.[5] Nevertheless, as

often known as turnover. As shown by Chernov et al. (2017), prepayments of out-the-money MBSs are mostly associated with turnover, while those of at-the-money and in-the-money MBSs are mostly associated with rate refinancing. The former type of prepayment behavior is more challenging to model, so prepayment forecast uncertainty and dealer dispersion is naturally larger for out-of-the-money cohorts than other cohorts.

[5]Disagreement can negatively affect asset returns in the presence of short-sale constraints (see Xiong (2013) and Hong and Stein (2007) for surveys of the literature

shown in Panel A of Table IA.XI, regressions of MBS yields on *Dealer Dispersion* alone deliver *negative* coefficients. Moreover, as reported in Panel B of Table IA.XI, the effects of *Dealer Dispersion* become insignificant once h^{WAOCS} is also included as a regressor, while the effects of h^{WAOCS} remain significant and positive. These findings suggest that the counter-intuitive negative regression coefficients of MBS yields on *Dealer Dispersion* mostly originate from the negative statistical correlation between *Dealer Dispersion* and our MBS heterogeneity measure h^{WAOCS}.

of asset pricing and heterogeneous beliefs). However, as discussed in Carlin et al. (2014), short-sale constraints are unlikely to be binding in MBS markets.

Table IA.I. MBS-Level Regressions of Prepayment Rates on Characteristics

	SMM^{6m}	SMM^{6m}	SMM^{6m}	SMM^{6m}	SMM^{12m}	SMM^{24m}	SMM^{36m}
WAOCS	0.024***			0.016***	0.028***	0.033***	0.024***
	(9.909)			(5.265)	(4.610)	(14.182)	(12.097)
WAOSIZE		0.012***		0.012***	0.023***	0.025***	0.023***
		(6.364)		(6.234)	(6.374)	(5.191)	(5.572)
WAOLTV			-0.012***	-0.013***	-0.024	0.008	0.049**
			(-4.256)	(-4.321)	(-1.395)	(0.688)	(2.911)
Intercept	-17.099***	-1.867***	1.268***	-12.720***	-22.374***	-30.335***	-20.980***
	(-9.775)	(-5.653)	(5.223)	(-6.575)	(-6.959)	(-12.318)	(-14.326)
Obs	97,591	97,591	97,591	97,591	92,520	80,985	67,104
R^2_{adj}	0.031	0.058	0.023	0.063	0.170	0.270	0.285
Coupon FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports results of individual-MBS-level regressions of average prepayment rates in the 6 months (SMM^{6m}), 12 months (SMM^{12m}), 24 months (SMM^{12m}), and 36 months (SMM^{12m}) since issuance on at-issuance characteristics, using newly-issued Fannie Mae 30-year TBA-eligible MBSs from January 2011 through December 2018. The coupon cohort fixed effects are included. The t-statistics based on robust standard errors clustered along the coupon dimension are reported in parentheses. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

Table IA.II. Summary Statistics of MBS Heterogeneity Measures

A: h^{WAOCS}						
Coupon	mean	min	p25	p50	p75	max
2.5	17.00	17	17	17	17	17
3	16.42	14	16	16	18	19
3.5	20.93	13	19	21	25	26
4	23.75	12	14	23	33	36
4.5	25.59	16	22	26	29	38
5	32.12	23	27	30	37	43
5.5	35.87	25	30	34	42	46
6	41.85	30	38	40	47	50
6.5	46.06	36	45	46	48	50
7	48.17	41	48	48	49	52
B: h^{SMM}						
Coupon	mean	min	p25	p50	p75	max
2.5	0.44	0.33	0.41	0.43	0.49	0.53
3	14.64	0.28	0.42	0.61	30.57	55.28
3.5	33.58	0.30	20.06	37.95	49.45	65.27
4	40.68	0.32	26.74	48.60	56.03	70.45
4.5	36.04	0.20	0.31	48.27	60.67	77.84
5	41.01	0.23	0.51	54.99	67.05	82.43
5.5	50.87	0.38	43.38	60.55	69.55	85.62
6	62.30	0.42	56.68	67.60	72.17	87.34
6.5	65.25	27.41	58.41	65.54	71.42	96.42
7	64.59	19.73	54.04	63.54	75.62	98.24

Note: In Panels A and B we report the summary statistics for monthly time-series of h^{WAOCS} and h^{SMM} including the quartiles for each coupon cohort of FNMA 30-year MBS.

Table IA.III. Summary of IO/PO Trusts

Trust	Collateral	Sample			Characteristics at Issuance					Time Series Means	
number	coupon	N	min	max	Amount ($bn)	Vintage	WAC	FICO	% LTV	moneyness	% SMM
405	4	19	201010	201204	2.63	2010	4.58	769	69	0.43	1.47
396	4.5	35	200906	201204	3.00	2009	4.94	759	68	0.62	1.35
400	4.5	30	200911	201204	2.00	2009	4.94	761	70	0.73	1.27
404	4.5	24	201005	201204	2.45	2010	4.96	761	71	0.88	1.42
337	5	86	200401	201102	1.55	2003	5.64	728	71	-0.16	1.32
340	5	100	200401	201204	2.24	2003	5.45	727	70	0.07	1.31
360	5	80	200507	201202	2.50	2005	5.69	728	71	0.11	1.23
377	5	66	200611	201204	3.77	2005	5.70	732	70	0.39	1.55
397	5	32	200909	201204	4.00	2009	5.49	745	76	1.20	1.97
346	5.5	98	200401	201202	2.00	2003	5.98	715	71	0.53	1.65
354	5.5	90	200411	201204	2.90	2004	5.94	715	73	0.62	1.56
363	5.5	75	200510	201202	2.05	2005	5.93	717	73	0.60	1.32
379	5.5	60	200705	201204	4.45	2007	6.10	723	74	1.00	2.09
399	5.5	30	200911	201204	2.15	2008	5.99	737	76	1.73	2.98
293	6	76	200401	201004	0.51	1994	6.70	715	72	0.66	1.46
344	6	76	200401	201004	2.20	2003	6.54	706	75	0.66	2.03
370	6	71	200606	201204	2.75	2006	6.43	711	75	1.28	1.77
372	6	55	200608	201102	3.00	2006	6.47	716	75	1.03	1.87
321	6.5	98	200401	201202	3.00	2002	6.99	699	75	1.53	2.22
371	6.5	69	200606	201202	2.50	2006	6.94	702	78	1.73	2.02
380	6.5	58	200707	201204	2.38	2006	7.06	699	80	2.05	2.29
320	7	86	200401	201102	3.03	2001	7.49	686	79	1.84	2.39
303	7.5	72	200401	200912	1.30	1999	7.95	693	78	2.11	2.22

Note: This table reports a summary of the monthly sample of IO/PO trusts. Trust numbers and collateral coupons are reported in the first two columns. The beginning month, the ending month, and the number of monthly observations are reported in the next three columns. The reported characteristics at the time of issuance include issuance amount (in $billion), vintage year, WAC, FICO score (weighted average across loans), and LTV ratios (in percentages). For each trust, the time-series mean of moneyness and SMM (in percentages) over the included sample period is also reported. The underlying collateral of all IO/PO trusts included are TBA-eligible FNMA 30-year Mega MBSs.

Table IA.IV. Likelihood of Equilibrium Types

Coupon	Obs	$\bar{c}$	$C^{\mathrm{SP}}_{95\%}$	$C^{\mathrm{SP}}_{90\%}$	$\Pr\{C^{\mathrm{SP}}_t < C^{\mathrm{SP}}_{95\%} - \bar{c}\}$	$\Pr\{C^{\mathrm{SP}}_t < C^{\mathrm{SP}}_{90\%} - \bar{c}\}$
3	923	14.31	34.86	27.25	81.15%	57.96%
3.5	856	14.21	34.83	27.13	81.66%	59.81%
4	814	17.01	47.86	31.73	88.82%	57.62%
4.5	785	18.78	48.63	36.56	82.80%	61.40%
5	736	30.39	88.55	60.24	88.99%	66.98%
5.5	790	30.43	100.69	66.39	90.63%	75.19%
6	803	39.48	114.46	83.95	88.17%	71.36%
6.5	879	67.47	226.90	160.63	89.76%	77.70%

Note: The first four columns report the number of days, the mean $\bar{c}$, the 95th percentile $C^{\mathrm{SP}}_{95\%}$, and the 90th percentile $C^{\mathrm{SP}}_{90\%}$ of the daily time series of SP trading costs, for each coupon cohort and from May 16, 2011 to August 5, 2015. The last two columns report the fraction of the days with the SP trading cost lower than the difference between the 95th percentile and mean and the difference between the 90th percentile and mean, respectively.

Table IA.V. Burnout

A: Regression of Prepayment Rates on Burnout		
Burnout	-0.066***	0.094
	(-5.429)	(1.217)
Burnout×Coupon		-0.026*
		(-1.904)
Intercept	0.476***	0.359***
	(27.414)	(6.791)
Obs	1,671	1,671
R^2_{adj}	0.073	0.077
Coupon FE	Yes	Yes

B: Regression of MBS Yields Controlling for Burnout						
	TBA-Eligible SP (LTV)				TBA Ineligible SP (LTV)	
	80-90	90-95	95-100	100-105	105-125	> 125
h^{WAOCS}	1.32***	1.33***	1.10***	0.84**	0.39	0.21
	(4.21)	(4.01)	(3.03)	(2.13)	(1.00)	(0.84)
SMM	-1.39***	-1.96***	-2.06***	-2.15***	-1.95***	-1.75**
	(-4.14)	(-5.84)	(-5.18)	(-5.71)	(-2.96)	(-2.55)
WAOLTV	-1.01	-0.77	-0.73	0.64	3.11	-0.84
	(-1.31)	(-0.70)	(-0.50)	(0.39)	(1.54)	(-1.25)
h^{Burnout}	0.58***	0.03	0.11	-0.18	0.47**	0.02
	(3.82)	(0.27)	(0.84)	(-1.47)	(2.26)	(0.09)
Intercept	84.21	82.95	89.18	1.52	-154.60	126.05***
	(1.52)	(1.05)	(0.86)	(0.01)	(-1.08)	(2.95)
Obs	390	390	390	390	390	390
R^2_{adj}	0.68	0.63	0.56	0.55	0.55	0.66
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Panel A reports regressions of SMM_{it} on Burnout_{it} and the interaction term between Burnout_{it} and coupon rate, controlling for the coupon cohort FE. Panel B reports the results for panel regressions of respective OASs of six groups of FNMA 30-year SP MBSs on h^{WAOCS}, with time FE. The t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. The overall sample period is from June 2003 through December 2018 in Panel A, and from June 2012 through December 2018 in Panel B. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

Table IA.VI. Alternative Sample

A: Number of CUSIPs in a Cohort ≥ 2000

	One-month Ahead		SP OAS						Liquidity Shock	SP/TBA Ratio	
	h_{t+1}^{SMM}	h_{t+1}^{WAOCS}	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP	Volume	# Trade
h_t^{WAOCS}	0.46***	0.99***	1.41***	1.33***	1.12***	0.83**	0.51	0.21	-1.27	0.11***	0.10***
	(2.63)	(475.65)	(5.16)	(4.10)	(3.20)	(2.08)	(1.39)	(0.84)	(-1.41)	(7.75)	(8.28)
SMM			-1.62***	-1.97***	-2.11***	-2.10***	-2.36***	-1.76***	-1.99***		
			(-5.13)	(-6.03)	(-5.50)	(-5.45)	(-3.54)	(-3.10)	(-3.28)		
WAOLTV			-1.21	-0.77	-0.77	0.64	3.02	-0.84	-0.90		
			(-1.20)	(-0.69)	(-0.51)	(0.42)	(1.32)	(-1.20)	(-1.10)		
$h_t^{\text{WAOCS}}\times$Distress									0.14**		
									(2.55)		
Distress									-3.21*		
									(-1.91)		
Issuance										0.02	-0.03***
										(1.64)	(-3.13)
Intercept	-12.08**	0.40***	99.26	83.06	92.02	1.21	-149.03	125.68***	145.17**	-6.21***	-3.83***
	(-2.58)	(6.43)	(1.41)	(1.04)	(0.85)	(0.01)	(-0.92)	(2.84)	(2.34)	(-19.99)	(-10.65)
Obs	1,411	1,411	390	390	390	390	390	390	2,340	368	368
R^2_{adj}	0.73	1.00	0.64	0.63	0.56	0.55	0.53	0.66	0.45	0.69	0.81
FE	Moneyness	Moneyness	Time	Time	Time	Time	Time	Time	SP Type	Time	Time

B: FHLMC 30-year MBS

	One-month Ahead		SP OAS						Liquidity Shock	SP/TBA Ratio	
	h_{t+1}^{SMM}	h_t^{WAOCS}	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP	Volume	# Trade
h_t^{WAOCS}	0.50**	0.99***	1.30***	1.04***	0.82***	0.43	0.36	0.47*	-1.50*	0.09***	0.05***
	(2.38)	(570.15)	(4.69)	(3.23)	(2.63)	(1.10)	(1.14)	(1.68)	(-1.70)	(5.41)	(3.50)
SMM			-1.32***	-1.53***	-1.59***	-1.56***	-2.12***	-1.57**	-1.71***		
			(-4.29)	(-4.87)	(-4.58)	(-4.08)	(-2.87)	(-2.37)	(-2.89)		
WAOLTV			-1.07	1.65	1.93	4.38	4.30	0.42	-0.82		
			(-0.45)	(0.62)	(0.62)	(1.36)	(1.16)	(0.28)	(-1.38)		
$h_t^{\text{WAOCS}}\times$Distress									0.16**		
									(2.47)		
Distress									-3.27*		
									(-1.83)		
Issuance										-0.04***	-0.05***
										(-3.22)	(-2.74)
Intercept	-12.39**	0.41***	97.37	-83.32	-94.90	-259.90	-243.70	34.00	142.25**	-4.02***	-1.13***
	(-2.32)	(18.46)	(0.60)	(-0.46)	(-0.44)	(-1.15)	(-0.94)	(0.31)	(2.51)	(-8.41)	(-3.50)
Obs	1,443	1,443	388	388	388	388	388	388	2,328	360	360
R^2_{adj}	0.70	1.00	0.55	0.55	0.50	0.51	0.52	0.64	0.41	0.55	0.57
FE	Moneyness	Moneyness	Time	Time	Time	Time	Time	Time	SP Type	Time	Time

Note: The first two columns report the results for panel regressions of one-month-ahead heterogeneity measures (h_{t+1}^{SMM} and h_{t+1}^{WAOCS}) on h_{it}^{WAOCS} using monthly data. The next six columns report results for panel regressions of the OAS of each of the six SP groups on h_t^{WAOCS}, controlling for SMM. The 9th column reports the panel regression of OASs on the interaction term $h_{it}^{\text{WAOCS}}\times$Distress by pooling all six SP groups together. The last two columns report results for panel regressions of the total monthly par volume (in \$billion) and number of trades on h_{it}^{WAOCS}, while controlling for monthly total new issuance (\$billion). In Panel A we limit the sample by excluding coupon cohorts with fewer than 2,000 outstanding MBSs, while for Panel B we use the alternative sample of FHLMC 30-year MBSs. Time fixed-effect are included in all but the regressions on the interaction term $h_{it}^{\text{WAOCS}}\times$Distress, where fixed effects for SP type are included. The t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

Table IA.VII. Measures of MBS Heterogeneity: the 10th Percentile and WAOSIZE

	A: 10th percentile										
	One-month Ahead		SP OAS						Liquidity Shock	SP/TBA Ratio	
	h_{t+1}^{SMM}	$h_{t+1}^{\text{WAOCS,10\%}}$	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP	Volume	# Trade
$h_t^{\text{WAOCS,10\%}}$	0.51*	0.99***	1.78***	1.67***	1.42***	1.04**	0.65	0.28	-1.54	0.14***	0.14***
	(1.91)	(421.30)	(4.93)	(3.87)	(3.05)	(1.97)	(1.38)	(0.85)	(-1.44)	(7.16)	(8.13)
SMM			-1.60***	-1.95***	-2.11***	-2.09***	-2.37***	-1.78***	-1.96***		
			(-5.02)	(-5.80)	(-5.33)	(-5.22)	(-3.56)	(-3.10)	(-3.18)		
WAOLTV			-2.23**	-1.72	-1.62	0.03	2.61	-1.02	-1.47		
			(-2.08)	(-1.62)	(-1.16)	(0.02)	(1.24)	(-1.30)	(-1.32)		
$h_t^{\text{WAOCS,10\%}}\times$Distress									0.18***		
									(2.61)		
Distress									-3.03**		
									(-1.98)		
Issuance										0.02	-0.03***
										(1.49)	(-3.34)
Intercept	-9.39*	0.23***	172.54**	150.93**	152.84	44.86	-119.70	138.69***	184.08**	-6.00***	-3.92***
	(-1.85)	(5.49)	(2.32)	(2.02)	(1.55)	(0.45)	(-0.81)	(2.80)	(2.37)	(-15.25)	(-12.14)
Obs	1,521	1,521	390	390	390	390	390	390	2,340	377	377
R^2_{adj}	0.74	1.00	0.64	0.62	0.56	0.54	0.53	0.66	0.45	0.67	0.81
FE	Moneyness	Moneyness	Time	Time	Time	Time	Time	Time	SP Type	Time	Time
	B: WAOSIZE										
	One-month Ahead		SP OAS						Liquidity Shock	SP/TBA Ratio	
	h_{t+1}^{SMM}	h_{t+1}^{WAOSIZE}	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP	Volume	# Trade
h_t^{WAOSIZE}	25.29**	0.77***	29.20***	14.97*	9.97	0.32	-2.93	-6.44	-33.62	3.71***	4.12***
	(2.28)	(6.97)	(3.51)	(1.93)	(1.30)	(0.04)	(-0.35)	(-0.87)	(-0.93)	(2.75)	(3.84)
SMM			-0.85***	-1.01***	-1.19***	-1.27***	-1.64**	-1.56***	-1.21***		
			(-3.58)	(-3.40)	(-3.48)	(-3.37)	(-2.23)	(-2.89)	(-4.85)		
WAOLTV			-2.76	0.45	0.91	3.25	4.41	0.14	-0.94		
			(-1.19)	(0.16)	(0.31)	(1.06)	(1.46)	(0.15)	(-1.03)		
$h_t^{\text{WAOSIZE}}\times$Distress									3.29		
									(1.36)		
Distress									-3.60		
									(-1.20)		
Issuance										-0.05***	-0.09***
										(-5.38)	(-11.51)
Intercept	-18.83**	0.15**	196.11	-1.70	-22.57	-172.30	-237.67	63.15	154.27**	-6.56***	-5.12***
	(-2.13)	(2.14)	(1.24)	(-0.01)	(-0.11)	(-0.82)	(-1.15)	(1.01)	(2.18)	(-4.74)	(-4.43)
Obs	1,519	1,517	390	390	390	390	390	390	2,340	377	377
R^2_{adj}	0.38	0.72	0.56	0.51	0.48	0.50	0.52	0.66	0.37	0.44	0.61
FE	Moneyness	Moneyness	Time	Time	Time	Time	Time	Time	SP Type	Time	Time

Note: The first two columns in Panel A report results for panel regressions using monthly data of one-month-ahead heterogeneity measures (h_{t+1}^{SMM} and $h_{t+1}^{\text{WAOCS,10\%}}$) on $h_{it}^{\text{WAOCS, 10\%}}$. The next six columns report results for panel regressions of the OAS of each of the six SP groups on $h_t^{\text{WAOCS,10\%}}$, controlling for SMM and WAOLTV. The 9th column reports the panel regression of OASs on the interaction term $h_{it}^{\text{WAOCS,10\%}}\times$Distress by pooling all six SP groups together. The last two columns report the results of panel regressions of the total monthly par volume (in \$billion) and number of trades on $h_{it}^{\text{WAOCS,10\%}}$, while controlling for monthly total new issuance (\$billion). In Panel B we use the h_{it}^{WAOSIZE} measure computed as the difference between the 95th percentile and the median of WAOSIZE. Time fixed-effects are included in all but the regressions on the interaction term $h_{it}^{\text{WAOCS}}\times$Distress, where fixed-effects for SP type are included. The t-statistics based on robust standard errors two-way that are clustered along the time and moneyness dimensions are reported in parentheses. Significance levels: *** for $p<0.01$, ** for $p<0.05$, and * for $p<0.1$, where p is the p-value.

Table IA.VIII. Measures of MBS Heterogeneity: Prepayment Rates and Combination of Different Characteristics

A: MBS heterogeneity based on prepayment rates									
	SP OAS						Liquidity Shock	SP/TBA Ratio	
	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP	Volume	# Trade
h_t^{SMM}	0.44***	0.40***	0.36***	0.31**	0.22**	-0.08	-0.81*	0.038***	0.039***
	(3.25)	(3.12)	(2.84)	(2.38)	(2.19)	(-0.96)	(-1.85)	(4.198)	(5.441)
SMM	-1.35***	-1.65***	-1.90***	-2.05***	-2.38***	-1.62***	-1.47***		
	(-4.38)	(-5.96)	(-6.19)	(-6.86)	(-3.50)	(-3.02)	(-2.82)		
WAOLTV	-1.82*	-1.37	-1.99**	-0.88	1.75	-1.71*	0.57		
	(-1.89)	(-1.49)	(-2.07)	(-1.39)	(0.80)	(-1.76)	(0.28)		
$h_t^{\text{SMM}}\times$Distress							0.06*		
							(1.75)		
Distress							-3.69*		
							(-1.82)		
Issuance								-0.013	-0.051***
								(-0.788)	(-4.140)
Intercept	163.88**	142.58*	200.66**	117.57**	-79.68	207.12***	61.52	-4.448***	-2.616***
	(2.18)	(1.89)	(2.49)	(2.22)	(-0.45)	(2.69)	(0.36)	(-9.719)	(-9.548)
Obs	390	390	390	390	390	390	2,340	377	377
R^2_{adj}	0.63	0.61	0.56	0.55	0.53	0.67	0.41	0.492	0.654
FE	Time	Time	Time	Time	Time	Time	SP Type	Time	Time

B: MBS heterogeneity measure combining different characteristics									
	SP OAS						Liquidity Shock	SP/TBA Ratio	
	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP	Volume	# Trade
h_t^{Combine}	1.11***	1.00***	0.84***	0.60**	0.33	0.11	-0.81	0.061***	0.059***
	(5.19)	(4.20)	(3.26)	(2.04)	(1.41)	(0.66)	(-1.47)	(5.941)	(7.284)
SMM	-1.58***	-1.93***	-2.09***	-2.07***	-2.27***	-1.80***	-1.85***		
	(-5.30)	(-6.13)	(-5.65)	(-5.58)	(-3.58)	(-3.40)	(-3.07)		
WAOLTV	-4.69***	-4.26***	-4.08***	-1.72	1.67	-2.30**	-0.96		
	(-3.95)	(-3.43)	(-2.65)	(-1.17)	(0.75)	(-2.10)	(-0.86)		
$h_t^{\text{Combine}}\times$Distress							0.09**		
							(2.29)		
Distress							-3.39*		
							(-1.87)		
Issuance								0.003	-0.039***
								(0.209)	(-4.413)
Intercept	382.83***	364.66***	360.59***	180.87	-74.90	253.01***	153.12*	-6.008***	-4.023***
	(4.09)	(3.62)	(2.90)	(1.56)	(-0.42)	(3.04)	(1.82)	(-9.906)	(-9.908)
Obs	390	390	390	390	390	390	2,340	377	377
R^2_{adj}	0.65	0.61	0.55	0.53	0.51	0.67	0.44	0.620	0.772
FE	Time	Time	Time	Time	Time	Time	SP Type	Time	Time

Note: The first six columns of Panel A report results for panel regressions of the OAS of each of the six SP groups on h_{it}^{SMM}. The 7th column reports the panel regression of OAS on the interaction term $h_{it}^{\text{SMM}}\times$Distress by pooling all six SP groups together. The last two columns report the results of panel regressions of the total monthly par volume (in \$billion) and number of trades on h_{it}^{SMM}, while controlling for monthly total new issuance (\$billion). For Panel B, we use the h_{it}^{Combine} measure computed as the fitted value of regressing $h_{i,t+1}^{\text{SMM}}$ on h_{it}^{WAOCS}, h_{it}^{WAOSIZE}, and h_{it}^{WAOLTV}. Time fixed-effects are included in all but the regressions on the interaction term with Distress, where fixed-effects for SP type are included. The t-statistics based on robust standard errors two-way that are clustered along the time and coupon dimensions are reported in parentheses. Significance levels: *** for $p<0.01$, ** for $p<0.05$, and * for $p<0.1$, where p is the p-value.

Table IA.IX. Alternative Measures of MBS Yields

	A: LIBOR OAS						
	80-90	90-95	95-100	100-105	105-125	> 125	ALL SP
h_t^{WAOCS}	1.54***	1.30***	1.19***	1.00***	1.00***	0.98***	-0.87
	(5.68)	(4.60)	(3.93)	(3.36)	(3.85)	(4.81)	(-1.00)
SMM	-1.83***	-1.72***	-1.78***	-1.71***	-2.06***	-1.78***	-1.61**
	(-6.53)	(-7.86)	(-7.81)	(-8.51)	(-4.97)	(-4.04)	(-2.24)
WAOLTV	-0.71	0.78	0.43	1.09	0.87	-1.73**	1.64
	(-0.56)	(0.45)	(0.22)	(0.70)	(0.59)	(-2.44)	(1.18)
$h_t^{\text{WAOCS}} \times$Distress							0.14***
							(3.84)
Distress							-2.62**
							(-2.00)
Intercept	62.68	-28.94	3.02	-36.35	-8.69	183.09***	-55.43
	(0.71)	(-0.24)	(0.02)	(-0.33)	(-0.08)	(4.24)	(-0.51)
Obs	390	390	390	390	390	390	2,340
R^2_{adj}	0.71	0.70	0.69	0.72	0.66	0.80	0.46
FE	Time	Time	Time	Time	Time	Time	SP Type
	B: OAS and Hedged Returns from another Dealer						
	OAS					Hedged Returns	
h_t^{WAOCS}	1.44***	1.42***				0.057***	0.032*
	(4.39)	(5.59)				(4.611)	(1.770)
SMM	0.08	-0.21				-0.012	0.022***
	(0.54)	(-0.99)				(-1.178)	(2.756)
WAOLTV	-1.55*	-1.40*				0.023	-0.069
	(-1.74)	(-1.70)				(0.562)	(-0.689)
$h_t^{\text{WAOCS}} \times$Distress		0.37***					0.016***
		(3.95)					(3.018)
Distress		-2.36					-0.776***
		(-0.64)					(-3.513)
Intercept	92.81	107.12*				-1.929	4.237
	(1.54)	(1.96)				(-0.659)	(0.601)
Obs	1,107	938				1,100	931
R^2_{adj}	0.82	0.30				0.746	0.090
FE	Time	No				Time	No

Note: The first six columns of Panel A report the results for panel regressions of respective OASs of six group of FNMA 30-year SP MBSs on h^{WAOCS}, while the last column reports the panel regression on the interaction term $h^{\text{WAOCS}} \times$Distress by pooling all six SP groups together. Time fixed-effects are included only in the first six columns, while the last column includes fixed-effects for SP type. In Panel B we report results of panel regressions of the OAS (in the first two columns) and of hedged returns (in the last two columns), obtained from an alternative major Wall Street MBS dealer, on h^{WAOCS} and the interaction term of $h^{\text{WAOCS}} \times$Distress separately. All regressions control for SMM and WAOLTV. The t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. The overall sample period is from June 2012 through December 2018 in Panel A and from June 2003 through December 2018 in Panel B. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

Table IA.X. Regressions of Trading Activity Controlling for Outstanding Balance

A: All Trades						
	Dollar Volume			Number of Trades		
	TBA	SP	SP/TBA	TBA	SP	SP/TBA
h^{WAOCS}	-7.645***	-0.483***	0.060***	-352.524***	-15.891	0.099***
	(-3.810)	(-3.112)	(6.229)	(-3.959)	(-1.216)	(12.417)
Outstanding	0.152	0.007	-0.005***	7.041	6.132***	-0.002*
	(1.284)	(0.865)	(-4.995)	(1.450)	(9.956)	(-1.806)
Intercept	360.217***	19.495**	-3.289***	14,429.979***	616.343	-3.533***
	(3.619)	(2.563)	(-12.367)	(3.625)	(0.949)	(-23.540)
Obs	377	377	377	377	377	377
R^2_{adj}	0.642	0.510	0.763	0.655	0.622	0.819
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

B: Dealer-Customer Trades						
	Dollar Volume			Number of Trades		
	TBA	SP	SP/TBA	TBA	SP	SP/TBA
h^{WAOCS}	-3.399***	-0.453***	0.052***	-67.384***	-18.587*	0.080***
	(-3.906)	(-3.258)	(5.289)	(-3.813)	(-1.882)	(10.299)
Outstanding	0.081*	0.004	-0.005***	2.317**	3.707***	-0.001
	(1.699)	(0.577)	(-4.425)	(2.514)	(7.606)	(-0.996)
Intercept	135.150***	17.673***	-2.468***	2,196.159***	646.539	-1.834***
	(3.537)	(2.606)	(-9.591)	(3.055)	(1.410)	(-11.346)
Obs	377	377	377	377	377	377
R^2_{adj}	0.623	0.503	0.708	0.599	0.592	0.747
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: In this table we report results of panel regressions of TBA and SP trading activities as well as their ratios on h^{WAOCS} for FNMA 30-year MBSs using monthly data. Trading activities are measured both by monthly total par volume (in \$billion) and by the total monthly number of trades. For Panel A we include all trades in computing measures of trading activities, while for Panel B we include only dealer-customer trades. All regressions control for monthly outstanding balance (in \$billion) and time fixed effects. t-statistics based on robust standard errors that are two-way clustered along the time and cohort dimensions are reported in parentheses. The overall sample period runs from June 2003 through December 2018 for TBA MBSs and from June 2012 through December 2018 for SP. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

Table IA.XI. Dealer Forecast Dispersion

	TBA-Eligible SP (LTV)				TBA Ineligible SP (LTV)		Dealer
	80-90	90-95	95-100	100-105	105-125	> 125	Dispersion
A: Regression on Dealer Dispersion							
SMM	-1.18***	-1.41***	-1.59***	-1.70***	-1.70***	-2.09***	0.24
	(-3.73)	(-3.93)	(-3.72)	(-4.01)	(-4.01)	(-2.84)	(1.33)
WAOLTV	-0.39	0.62	0.52	1.54	3.11	0.42	3.69**
	(-0.25)	(0.32)	(0.24)	(0.70)	(1.30)	(0.38)	(2.30)
Dealer Dispersion	-17.20***	-14.09***	-11.88***	-9.16*	-7.06	-11.66**	
	(-4.90)	(-3.39)	(-2.59)	(-1.94)	(-1.39)	(-2.12)	
CP							-0.54***
							(-9.30)
Intercept	77.63	17.57	28.52	-40.92	-140.59	70.44	3.11***
	(0.70)	(0.13)	(0.19)	(-0.27)	(-0.87)	(0.97)	(12.99)
Obs	390	390	390	390	390	390	390
R^2_{adj}	0.62	0.58	0.53	0.53	0.54	0.53	0.70
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B: Regression on h^{WAOCS} Controlling for Dealer Dispersion							
h^{WAOCS}	0.97***	1.06**	0.89*	0.64	0.21	0.28	
	(3.39)	(2.49)	(1.86)	(1.02)	(0.48)	(1.24)	
SMM	-1.63***	-1.97***	-2.10***	-2.08***	-2.26***	-1.81***	
	(-5.24)	(-6.05)	(-5.35)	(-5.35)	(-3.19)	(-3.01)	
WAOLTV	-1.62	-1.01	-0.91	0.54	2.95	-0.83	
	(-1.43)	(-0.85)	(-0.57)	(0.34)	(1.27)	(-1.26)	
Dealer Dispersion	-9.17**	-5.53	-4.46	-3.69	-5.23	1.64	
	(-2.56)	(-1.04)	(-0.70)	(-0.48)	(-0.82)	(0.33)	
Intercept	144.17*	109.93	109.64	14.86	-134.32	122.29***	
	(1.83)	(1.31)	(0.98)	(0.13)	(-0.84)	(3.29)	
Obs	390	390	390	390	390	390	
R^2_{adj}	0.67	0.63	0.57	0.55	0.54	0.66	
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	

Note: In the first six columns of Panel A we report the results for panel regressions of respective OASs of six groups of FNMA 30-year SP MBSs on Dealer Dispersion, while in the last column we report the panel regression of Dealer Dispersion on coupon rate (CP). In Panel B we report results of panel regressions of the OAS on h^{WAOCS} controlling for Dealer Dispersion. All regressions control for SMM, WAOLTV, and time fixed effects. The t-statistics based on robust standard errors that are two-way clustered along the time and moneyness cohort dimensions are reported in parentheses. The sample period is from June 2012 through December 2018. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

REFERENCES

Belikoff, Alexander, Kirill Levin, Harvey Stein, and Xusheng Tian, 2010, Analysis of mortgage-backed securities: Before and after the credit crisis, *Bloomberg LP, Verson 2.2* .

Carlin, Bruce, Francis Longstaff, and Kyle Matoba, 2014, Disagreement and asset prices, *Journal of Financial Economics* 114, 226–238.

Chernov, Mikhail, Brett R. Dunn, and Francis A. Longstaff, 2017, Macroeconomic-Driven Prepayment Risk and the Valuation of Mortgage-Backed Securities, *The Review of Financial Studies* 31, 1132–1183.

Diep, Peter, Andrea L. Eisfeldt, and Scott Richardson, 2021, The cross section of mbs returns, *The Journal of Finance* forthcoming.

Fabozzi, Frank, and Steven Mann, 2011, Handbook of fixed income securities, *McGraw-Hill; 8 edition* .

Hendershott, Terrence, and Ananth Madhavan, 2015, Click or call? auction versus search in the over-the-counter market, *The Journal of Finance* 70, 419–447.

Himmelberg, Charles, Marty Young, Hui Shan, and Chris Henson, 2013, The impact of fed mbs purchases on tba pool quality, *Goldman Sachs* .

Hong, Harrison, and Jeremy C. Stein, 2007, Disagreement and the stock market, *The Journal of Economic Perspectives* 21, 109–128.

Schwartz, Eduardo S., and Walter N. Torous, 1993, Mortgage prepayment and default decisions: A poisson regression approach, *Real Estate Economics* 21, 431–449.

Song, Zhaogang, and Haoxiang Zhu, 2019, Mortgage Dollar Roll, *The Review of Financial Studies* 32, 2955–2996.

Xiong, Wei, 2013, Bubbles, crises, and heterogeneous beliefs, *Handbook for Systemic Risk* 663–713.