

Asset Pricing with Cohort-Based Trading in MBS Markets *

Nicola Fusari,[†] Wei Li,[‡] Haoyang Liu,[§] Zhaogang Song[¶]

Last updated: November 29, 2020

Abstract

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

Keywords: Cohort, Heterogeneity, Liquidity, MBS, Prepayment, TBA

JEL classification: G12, G18, G21, E58

*We are grateful to Doug McManus at Freddie Mac and Marty Young at Goldman Sachs for helpful conversations on the MBS market. For helpful comments, we thank 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 views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

[†]The Johns Hopkins Carey Business School; Email: nfusari1@jhu.edu.

[‡]The Johns Hopkins Carey Business School; Email: wei.li@jhu.edu.

[§]Federal Reserve Bank of New York, E-mail: haoyang.liu@ny.frb.org.

[¶]The Johns Hopkins Carey Business School; Email: zsong8@jhu.edu.

1 Introduction

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). The agency MBS market has 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, Logan, Martin, McCabe, Natalucci and Remache (2015); FRBNY (2020)). Agency MBSs are also among the most important assets on the balance sheets of large financial institutions such as insurance companies and banks (Ihrig, Kim, Vojtech and Weinbach (2019); Chodorow-Reich, Haddad and Ghent (2019)).

Despite the importance of the agency MBS market, only a few studies have examined variations in MBS returns, most of which focus on prepayment risks associated with the uncertain timing of cash flows.¹ Differing from these studies, we investigate how the unique agency MBS market microstructure influences MBS returns.² 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 MBSs are traded at the same price using a standardized contract. A TBA contract specifies, for example, only that a delivered MBS is guaranteed by Fannie Mae, consists of 30-year fixed-rate mortgages, and pays a coupon of 4%, usually known as a coupon cohort. We show in this paper that this unique parallel trading environment influences MBS pricing and trading through distinctive economic channels, resulting in large return variations on top of those driven by prepayment risk.

To guide our empirical analysis, we propose a simple model to demonstrate the economic channels through which the unique agency MBS market structure affects the trading and pricing of 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

¹The uncertain timing of cash flows arises because mortgage borrowers can prepay loan balances without penalty, and would do so particularly when interest rates are low. See, for example, the recent contributions by Gabaix, Krishnamurthy and Vigneron (2007), Chernov, Dunn and Longstaff (2017), Boyarchenko, Fuster and Lucca (2019), and Diep, Eisfeldt and Richardson (2019).

²See Easley and O'Hara (2003), Amihud, Mendelson and Pedersen (2006), and Vayanos and Wang (2013) for broad surveys of studies of market microstructure, liquidity, and asset pricing.

when choosing between TBA and SP trading. On the one hand, sellers incur higher transaction costs 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, Maxwell and Venkataraman (2013) and Gao, Schultz and Song (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 eligible MBSs, resulting in a CTD price discount to MBSs traded on the TBA market. Such a discount is absent from the SP market because every MBS is priced individually.

Two distinctive economic channels arise from this parallel trading environment. First, high-value MBSs are more likely to be sold in the SP market. Intuitively, because a single TBA price is set for any eligible MBS, sellers of higher-value MBSs have to accept greater CTD discounts when they use the TBA market.³ We call this the *venue selection* channel. Second, round-1 buyers in the SP market can use TBA as a backup selling venue in round 2 when they experience liquidity shocks. 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" (Gao et al. (2017)) in the future. We call this 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, greater MBS heterogeneity increases the CTD discount, making more sellers choose the SP market. Second, a greater CTD discount reduces the value of the round-2 TBA market as a backup selling venue, so SP buyers in round 1 would lower their bid prices.

To empirically measure MBS heterogeneity, we use the difference in prepayment characteristics between the cheapest and average MBSs within a coupon cohort. Specifically, we obtain monthly series of weighted-average original FICO scores (WAOCS), a key input to most MBS prepayment models, for all individual Fannie Mae 30-year MBSs between June 2003 and December 2018. An increase in WAOCS is associated with higher prepayment risk and lower MBS value (Fabozzi and Mann (2011)). 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.⁴ Regressions of h^{WAOCS} on lagged h^{WAOCS} de-

³Downing, Jaffee and Wallace (2009) show that MBSs backing up collateralized mortgage obligation deals are of lower quality than others, similar to the venue selection between the TBA and SP markets.

⁴All our main results remain robust to using other relevant MBS characteristics, for example the weighted aver-

liver positive and highly significant coefficients, showing that this variable captures *expected* MBS heterogeneity in the cross section reasonably well.

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

First, although the CTD issue is absent from SP trading (and present in TBA contracts) by design, MBS heterogeneity associated with the CTD discount affects returns of SP MBSs positively through the venue backup channel. When MBS heterogeneity increases, the TBA market as a future backup selling venue is less valuable to today's SP buyers, who then demand a higher return as compensation. Further, because of the venue selection channel, the effect of MBS heterogeneity on SP returns is weaker for higher-value SP MBSs because they are less likely to be sold on the TBA market in the future. These effects of MBS heterogeneity on SP returns 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.⁵

We hence focus on testing the effects of MBS heterogeneity on SP returns in our main analyses. We follow [Gabaix et al. \(2007\)](#), [Boyarchenko et al. \(2019\)](#), and [Song and Zhu \(2019\)](#) to measure MBS returns with the option-adjusted spread (OAS). As [Boyarchenko et al. \(2019\)](#) show, the OAS contains returns caused by prepayment risk and market illiquidity. Thus, our analysis relates the market illiquidity component to the TBA/SP parallel trading environment. Within each coupon cohort, we consider SP MBSs with distinct loan-to-value (LTV) ratios.⁶ Consistent with our hypotheses, for SP MBSs with loan-to-value (LTV) ratios in the 80%-90% range, which are very 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 16 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, the associated increase in the OAS is about 10 basis points. In addition, we show that h^{WAOCS} affects the OAS of TBA MBSs, consistent with the presence of CTD discount.

Second, the effects of MBS heterogeneity on MBS returns are amplified when future liquid-

age original loan size (WAOSIZE), or to using other percentiles (e.g. the 90th percentile). To avoid potential outliers, we do not use the 100th percentile. See [Section 7](#) for details.

⁵Early studies have examined the CTD discounts in futures contracts, including e.g. [Hegde \(1988\)](#), [Hemler \(1990\)](#), [Kane and Marcus \(1986\)](#), and [Gay and Manaster \(1984\)](#), among others.

⁶Using 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.

ity shocks are more likely to occur. As our model shows, when the likelihood of future liquidity shocks increases, traders are more likely to use TBA markets in the future, making SP returns more sensitive to MBS heterogeneity. We use the VIX index to proxy for selling pressure, which is used in many studies as a measure of broad market stress. We also supplement the VIX index with the Distress measure of [He, Khorrami and Song \(2019\)](#),⁷ 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 VIX and Distress generate positive and highly significant coefficients, confirming the stronger effects of MBS heterogeneity on SP returns 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 thus reduces the liquidity benefits of TBA trading per se and as a backup venue. This should dampen trading activities on both the SP and TBA markets. Further, the increased CTD price discount would make sellers less willing to use the TBA market relative to the SP market, thus 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 of about \$62 billion and \$4 billion in TBA and SP monthly trading volume, respectively, as well as a decrease of about 138% in the ratio of SP volume to TBA volume.

We conduct numerous robustness checks. Our main results remain significant after controlling for potential prepayment model misspecification underlying the OAS measures, using alternative measures of MBS heterogeneity, using alternative samples, using OAS based on the Libor swap curve, and so on. Importantly, we perform two analyses that distinguish the effects of MBS heterogeneity from that of the prepayment risks. The first analysis hinges on the findings of [Gabaix et al. \(2007\)](#) and [Diep et al. \(2019\)](#) that the market price of prepayment risk shows opposite signs depending on whether the aggregate MBS market is dominated by premium or discount securities. We find that the impact of MBS heterogeneity is positive regardless of whether premium or discount securities dominate the aggregate MBS market. The second

⁷The 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, Kelly and Manela \(2017\)](#) and the market-price-based “noise” measure of [Hu, Pan and Wang \(2013\)](#).

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 returns of both IO and PO strips, which goes against interpreting our heterogeneity measure as a proxy for prepayment risks.

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 volume of TBA contracts arises from investors' hedging and speculation activities that are often reversed before maturity and do not lead to deliveries of MBSs. In fact, a rough estimate in [An, Li and Song \(2020\)](#) shows that slightly more than half of the newly issued TBA-eligible MBSs are 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 3](#)).⁸

Overall, we find fundamental and large economic impacts of the unique MBS market structure with parallel TBA and SP trading. These results are of broad interest because parallel trading venues are present in many markets, e.g. dark pool and exchange trading in equity markets, futures and cash trading in Treasury markets, etc. Furthermore, TBA-like trading mechanisms have been advocated for other fixed-income markets such as corporate bonds and municipal bonds. For example, [Bessembinder, Spatt and Venkataraman \(2019\)](#) 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." The economic channels we document shed light on the potential effects of introducing such a market design.

Related Literature. Our paper contributes to the asset pricing literature on MBS markets, most studies in which focus on prepayment risks. Early studies proposed valuation frame-

⁸In addition, the TBA-ineligible MBSs excluded from our analysis, usually pools backed by high-balance mortgages, forty-year mortgages, interest-only mortgages, and so on, only account for a small fraction of agency MBSs (less than 1% according to [An, Li and Song \(2020\)](#)).

works based on the contingent claims approach and econometric prepayment models.⁹ Recently, [Levin and 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. \(2019\)](#) study the prepayment risk premium under a limits-of-arbitrage framework, while [Duarte, Longstaff and Yu \(2007\)](#) document characteristics of various MBS portfolio strategies.¹⁰

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 \(2019\)](#) present evidence on the existence of scarcity premium and convenience premium for agency MBSs, while [Bartolini, Hilton, Sundaresan and Tonetti \(2011\)](#) and [Song and Zhu \(2019\)](#) examine the premium of MBS as collateral in funding markets. Differing from these papers, ours shows that the unique parallel trading environment affects MBS returns 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, Jankowitsch and Subrahmanyam \(2017\)](#), [Gao, Schultz and Song \(2018\)](#), [Schultz and Song \(2019\)](#), and [Kim and Huh \(2019\)](#). 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 \(2013\)](#).

2 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

⁹The 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, Richardson, Stanton and Whitelaw \(1997\)](#), [Titman and Torous \(1989\)](#), [Downing, Stanton and Wallace \(2005\)](#), and [Longstaff \(2005\)](#) among other studies. Studies based on econometric prepayment models and focused on analysis of OAS include [Schwartz and Torous \(1989\)](#), [Richard and Roll \(1989\)](#), and [Deng, Quigley and Van Order \(2000\)](#). The prepayment model framework has been extended by [Chen \(1996\)](#) and [Cheyette \(1996\)](#) to estimate implied prepayments from MBS prices.

¹⁰Relatedly, [Duarte \(2007\)](#), [Malkhozov, Mueller, Vedolin and Venter \(2016\)](#) and [Hansen \(2014\)](#) study the effects of mortgage-risk hedging on Treasury and broader interest rate markets.

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.¹¹ Our analyses focus mainly on pass-through MBSs, but do 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 drop, 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. There is substantial heterogeneity in prepayment risk across individual MBSs because each MBS is “unique in its prepayment characteristics” (Gao et al. (2017)). This arises from the vastly different characteristics of mortgage loans and their borrowers (see Section 4 for summary statistics of different prepayment speeds of varying MBSs).

One might think, given the large asset heterogeneity and OTC nature of trading, that the agency MBS market would be very illiquid, just like the illiquid corporate and municipal bond markets (Bessembinder et al. (2019)). On the contrary, a large portion of agency MBSs are traded at low transaction costs of only about 2 basis points in the TBA market, 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.¹² 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.

However, the single cohort-level price for heterogeneous MBSs, leads naturally to a CTD issue and price discount on TBA MBSs. Intuitively, the TBA price discount relates positively to the

¹¹According to SIFMA, the outstanding balances of pass-through and structured MBSs are about \$7.3 and \$1.1 trillion, respectively.

¹²There is one settlement day per month, set by the SIFMA. It also sets the eligibility criterion for TBA delivery, known as the “Good Delivery” requirement. Details on the TBA settlement schedule are available at <http://bit.do/sifma-mbsset>, while those on the TBA eligibility criterion can be found at <http://bit.do/sifma-umbs8>.

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)). Instead, TBA-eligible MBSs can be traded in either the TBA or 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” to TBA prices. However, SP trading incurs high transaction costs of about 20-60 basis points. 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. The liquidity value TBA trading provides to SP trading comes 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.

3 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. Guided by the model, we set up the hypotheses for empirical testing.

3.1 A Simple Model of MBS Trading and Pricing

We abstract prepayment risk away from the model to focus on how the parallel trading environment affects the trading and pricing of MBSs in the TBA and SP markets. In particular, we consider MBSs that are eligible for trading in both the TBA market and the SP market. The time discount rate is normalized at zero. MBSs are issued at time 1 and mature at time 3. At time 2,

a fraction ρ of MBS owners experience idiosyncratic liquidity shocks, forcing them to sell their MBSs. Hence 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 receiving its terminal payoff at time 3. The time-3 payoff of an MBS v follows a distribution with density $f(\cdot)$ over the range $[v_m - h_d, v_m + h_u]$, where v_m is the median MBS payoff and is assumed to be fixed. The parameter h_d , the difference in value between the median and the cheapest MBSs, conveniently captures the cross-sectional dispersion of MBS values, among many other measures. We term it as (downside) MBS heterogeneity.¹³

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 the uniform price P_t^{TBA} at time $t \in \{1, 2\}$. This is the CTD issue in the TBA market, which is similar to the “lemon’s problem” described by Akerlof (1970). We assume, for simplicity, that TBA buyers recognize the CTD issue and bid

$$P_t^{\text{TBA}} = v_m - h_d. \quad (1)$$

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 $v_k - P_t^{\text{TBA}} = v_k - v_m + h_d$, which increases with MBS heterogeneity h_d (relative to the fixed v_m). Our main results hold with a more general TBA pricing $P_t^{\text{TBA}} = g(h_d; v_m)$ that decreases with h_d for a 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 to pay a cost C_t^{SP} to locate a buyer. At time 1, before buyers bid and sellers choose the selling venue, they observe the current transaction cost C_1^{SP} and believe that C_2^{SP} , the future transaction cost at time 2, follows a two-point distribution,

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

¹³The parameter h_u measures the upside MBS heterogeneity. It is less important because high-value MBSs are sold on SP markets in which cross-sectional heterogeneity does not affect pricing.

for $c_{2,h} \geq c_{2,\ell} \geq 0$. At time 2, sellers choose the selling venue after observing C_2^{SP} .¹⁴

We find the equilibrium using backward induction. We assume for simplicity that buyers in the SP market at time 2 earn zero profits and bid the fair value

$$P_2^{\text{SP}}(v_k) = v_k \quad (3)$$

for MBS k . Then the equilibrium at time 2 is straightforward because every MBS will mature and pays its fundamental value at time 3.

Proposition 1 (Time 2 equilibrium). *At time 2, a seller sells her MBS on the TBA market at price $P_2^{\text{TBA}} = v_m - h_d$ if the value of her MBS $v \leq \bar{v}_2$ and on the SP market at price $P_2^{\text{SP}}(v) = v$ if the value of her MBS $v > \bar{v}_2$, where the TBA value threshold is*

$$\bar{v}_2 := P_2^{\text{TBA}} + C_2^{\text{SP}} = v_m - h_d + C_2^{\text{SP}}. \quad (4)$$

For an MBS whose value $v > \bar{v}_2$, its CTD price discount in the TBA market $v - P_2^{\text{TBA}}$ is greater than its explicit SP selling cost C_2^{SP} . Hence its seller chooses the SP market. MBSs whose values $v \leq \bar{v}_2$ are sold on the TBA market because the CTD discount $v - P_2^{\text{TBA}} \leq C_2^{\text{SP}}$.

Ascertaining the equilibrium SP price at time 1 is less straightforward. Because the owner of an MBS might be forced to sell it at time 2, an SP buyer at time 1 would bid a price that is equal to an MBS's terminal payoff less its expected effective selling cost at time 2. The complication is that this effective selling cost depends on which market (i.e. the TBA or SP) is used.

According to [Proposition 1](#), because the future selling cost C_2^{SP} is random, the time-2 TBA value threshold follows the following distribution

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

where

$$\bar{v}_{2,h} := P_2^{\text{TBA}} + c_{2,h} = v_m - h_d + c_{2,h} \quad \text{and} \quad \bar{v}_{2,\ell} := P_2^{\text{TBA}} + c_{2,\ell} = v_m - h_d + c_{2,\ell}. \quad (6)$$

¹⁴We can think of π_h as the risk-neutral probability of $C_2^{\text{SP}} = c_{2,h}$. The randomness of SP transaction costs are documented in empirical studies, for example [Gao et al. \(2017\)](#). If C_2^{SP} is deterministic, we can show that SP buyers of time 1 would never sell in the TBA market at time 2 and MBS heterogeneity has no impact on SP returns.

If an investor is forced to sell at time 2 because of liquidity shocks, the effective selling cost depends on the value of the MBS. Specifically, because a high-value MBS ($v_k > \bar{v}_{2,h}$) is always sold through the SP market, its effective selling cost is C_2^{SP} . Because a low-value MBS ($v_k < \bar{v}_{2,\ell}$) is always sold through the TBA market, its effective selling cost is $v_k - P_2^{\text{TBA}} = v_k - v_m + h_d$. Because a medium-value MBS ($\bar{v}_{2,\ell} \leq v_k \leq \bar{v}_{2,h}$) will be sold at time 2 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, its expected effective selling cost 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*

$$P_1^{\text{SP}}(v_k) = \begin{cases} v_k - \rho \mathbf{E}[C_2^{\text{SP}}] & \text{if } v_k > \bar{v}_{2,h}, \\ v_k - \rho [\pi_h(v_k - v_m + h_d) + (1 - \pi_h)c_{2,\ell}] & \text{if } \bar{v}_{2,\ell} \leq v_k \leq \bar{v}_{2,h}, \\ v_k - \rho(v_k - v_m + h_d) & \text{if } v_k < \bar{v}_{2,\ell} \end{cases} \quad (7)$$

for an asset of value v_k .

Figure 1 illustrates the impact of having a TBA market at time 2 on the time-1 SP price. Without a TBA market at time 2, any MBS is sold on the SP market at time 2 and thus $P_1^{\text{SP}}(v_k) = v_k - \rho \mathbf{E}[C_2^{\text{SP}}]$ for all v_k (the red dashed line). The existence of the TBA market at time 2 lowers the expected selling cost at time 2. The existence of TBA market raises an asset's time-1 SP price (the blue solid line) if its value $v \leq \bar{v}_{2,h}$.

We now describe the equilibrium at $t = 1$. If the seller of MBS k chooses the SP market, she spends C_1^{SP} to find a buyer and receives $P_1^{\text{SP}}(v_k)$ from that buyer, realizing a net revenue of $P_1^{\text{SP}}(v_k) - C_1^{\text{SP}}$. If the seller chooses the TBA market, she receives P_1^{TBA} . To maximize her revenue, the seller chooses the TBA market if $P_1^{\text{TBA}} > P_1^{\text{SP}}(v_k) - C_1^{\text{SP}}$ and the SP market otherwise. Naturally, the time-1 TBA threshold \bar{v}_1 will be the asset value that equates the revenues from the two markets.

Depending on parameter values, the time-1 TBA threshold \bar{v}_1 may be greater than $\bar{v}_{2,h}$, less than $\bar{v}_{2,\ell}$, or fall between $\bar{v}_{2,\ell}$ and $\bar{v}_{2,h}$, resulting in three types of equilibrium. Here we focus on the last equilibrium type with $\bar{v}_{2,\ell} < \bar{v}_1 < \bar{v}_{2,h}$ to demonstrate the effects of future TBA trading on current SP prices conveniently. Appendix A describes the other two types of equilibria.

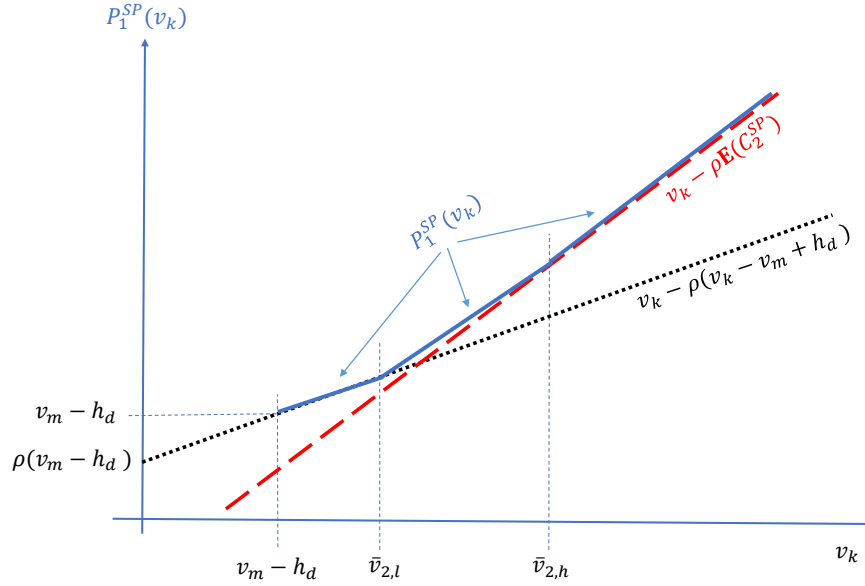


Figure 1. Time-1 SP price $P_1^{\text{SP}}(v_k)$.

Proposition 2 (Time 1 equilibrium). When $(1 - \rho)c_{2,\ell} \leq C_1^{\text{SP}} \leq c_{2,h} - \rho \mathbf{E}[C_2^{\text{SP}}] = (1 - \rho\pi_h)(c_{2,h} - c_{2,\ell}) + (1 - \rho)c_{2,\ell}$, the TBA threshold at time 1 is

$$\bar{v}_1 = v_m - h_d + \frac{C_1^{\text{SP}} + \rho(1 - \pi_h)c_{2,\ell}}{1 - \rho\pi_h}, \quad (8)$$

which falls within $[\bar{v}_{2,\ell}, \bar{v}_{2,h}]$. At time 1, the MBS of value v_k is sold on the TBA market at price $P_1^{\text{TBA}} = v_m - h_d$ if $v_k < \bar{v}_1$ and sold on the SP market at price

$$P_1^{\text{SP}}(v_k) = \begin{cases} v_k - \rho [\pi_h(v_k - v_m + h_d) + (1 - \pi_h)c_{2,\ell}] & \text{if } \bar{v}_1 \leq v_k \leq \bar{v}_{2,h}, \\ v_k - \rho \mathbf{E}[C_2^{\text{SP}}] & \text{if } v_k > \bar{v}_{2,h}, \end{cases} \quad (9)$$

if $v_k \geq \bar{v}_1$.

That is, low-value MBSs ($v_k < \bar{v}_1$) are sold on the TBA market at price P_1^{TBA} at time 1. Both medium-value ($\bar{v}_1 \leq v_k \leq \bar{v}_{2,h}$) and high-value ($v_k > \bar{v}_{2,h}$) MBSs are sold on the SP market at time 1. The discount $v_k - P_1^{\text{SP}}(v_k)$ reflects an SP buyer's expected effective selling cost at time 2,

which differs for medium-value and high-value MBSs because medium-value MBSs might be sold on the TBA market at time 2 while high-value MBSs are always sold on the SP market.

Overall, the key insight from the model is that, because owners of SP MBSs may use the TBA market as a backup selling venue when they experience liquidity shocks, the magnitude of the CTD price discount in the TBA market can influence the prices and returns 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 CTD discount has on the price of this SP MBS today.

Before moving on to develop testing hypotheses based on the model, we provide a few discussions on the model setup.

First, generally speaking, we study the impact of transaction costs on asset returns 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 the CTD issue in the TBA market can influence MBS returns in the SP market. That is, the very existence of the TBA market can reduce expected *future* transaction costs for certain SP assets.

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 returns—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 (see [Li and Song \(2020\)](#) for a search-based theoretical model along this direction). 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.

Third, as mentioned above, the equilibrium characterized in [Proposition 2](#) is one of the three equilibria. In this equilibrium, certain time-1 SP MBSs have less than 100% but positive chances of being sold on the TBA market at time 2. We believe this is likely the case in practice, corresponding to the description of [Gao et al. \(2017\)](#) that TBA trading gives “potential buyers of an SP an option to deliver the SP in a TBA trade if market conditions change.” The other two equilibria are presented in [Appendix A](#).

3.2 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. We conduct comparative statics analyses by varying h_d , given a fixed v_m .

When h_d increases, TBA buyers expect sellers to deliver worse MBSs and lower their bid prices accordingly, resulting in a deeper price discount $v_k - P_t^{\text{TBA}} = v_k - v_m + h_d$ for an MBS with value v_k if it is traded in the TBA market. Such a discount is specific to the TBA market (and in fact, is present in all contracts with CTD features, e.g. Treasury futures) and does not depend on 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 reflects the impact of the parallel trading environment. In particular, [Proposition 2](#) implies that the realized return (or yield) of an MBS sold on the SP market at time 1 is

$$y_1^{\text{SP}}(v_k) = \frac{v_k}{P_1^{\text{SP}}(v_k)} - 1 = \begin{cases} \frac{v_k}{v_k - \rho[\pi_h(v_k - v_m + h_d) + (1 - \pi_h)c_{2,\ell}]} - 1 & \text{if } \bar{v}_1 \leq v_k \leq \bar{v}_{2,h}, \\ \frac{v_k}{v_k - \rho E[C_2^{\text{SP}}]} - 1 & \text{if } v_k > \bar{v}_{2,h} \end{cases} \quad (10)$$

Although SP trading does not involve any CTD issue by design, the returns of medium-value SP MBSs ($v_k \in [\bar{v}_1, \bar{v}_{2,h}]$) are affected by MBS heterogeneity h_d because these MBSs might be sold through the TBA market at price $P_2^{\text{TBA}} = v_m - h_d$ in the future at time 2. Therefore, the parallel trading environment establishes a channel through which MBS heterogeneity can affect the return of MBSs traded in the SP market.

To see the direction of the effect more clearly, we examine the impact of h_d on $\frac{v_k - P_1^{\text{SP}}(v_k)}{v_k} = \frac{y_1^{\text{SP}}(v_k)}{1 + y_1^{\text{SP}}(v_k)}$, a monotonically increasing transformation of $y_1^{\text{SP}}(v_k)$ that is easier to analyze. For a medium-value MBS sold in the SP market at time-1, the marginal impact

$$\frac{\partial}{\partial h_d} \left(\frac{y_1^{\text{SP}}(v_k)}{1 + y_1^{\text{SP}}(v_k)} \right) = \frac{\rho \pi_h}{v_k} \quad (11)$$

is positive. Hence, the returns of medium-value SP increase with MBS heterogeneity h_d . Intuitively, when the MBS cohort is more heterogeneous, the TBA price falls, which diminishes the value of the future TBA market as a backup selling venue for these medium-value SP MBSs. Consequently, buyers lower bid prices for these SP MBSs to compensate for the drop in poten-

tial resale revenue. Further, the effect of MBS heterogeneity on SP returns is not uniform across MBSs. Specifically, Eq. (10) shows that the returns of high-value MBSs ($v_k > \bar{v}_{2,h}$) do not depend on h_d because these MBSs will never be sold on the TBA markets. Moreover, for medium-value SP MBSs, Eq. (11) shows that the magnitude of the dependence ($\frac{\rho\pi_h}{v_k}$) decreases with v_k . Intuitively, because a more valuable MBS is less likely to be sold on the TBA market, its return is 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 increases ($h_d \uparrow$), the yield of an MBS traded on the SP market increases ($y_1^{\text{SP}}(v_k) \uparrow$), and the effect is weaker for a more valuable SP MBS ($v_k \uparrow$).*

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. For a medium-value SP MBS ($v_k \in [\bar{v}_1, \bar{v}_{2,h}]$), when ρ increases, Eq. (11) shows that the dependence of the yield on MBS heterogeneity is stronger (a larger $\frac{\rho\pi_h}{v_k}$). Intuitively, TBA trading as a backup selling venue is more important when an SP buyer is more likely to experience a liquidity shock 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 Eq. (8) shows, an increase in the MBS heterogeneity h_d leads to a lower \bar{v}_1 , the upper bound of MBS value traded in the TBA market. 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 $v_k - P_1^{\text{TBA}} = v_k - v_m + h_d$ and pushing sellers towards the SP market. We state the hypothesis as follows.

Hypothesis 3.1. *When MBS heterogeneity increases, the proportion of MBSs traded on the SP market increases.*

Moreover, given that the existence of the TBA market reduces transaction costs across the entire MBS market, it is reasonable to expect that, when MBS heterogeneity h_d increases, the TBA price discount is greater and the cost-saving benefit of TBA trading diminishes for any

MBS, thereby reducing trading activities across the entire market.¹⁵ We formulate this hypothesis as follows.

Hypothesis 3.2. *When MBS heterogeneity increases, trading activities decline 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, based on [Proposition 2](#), our model implies that the realized yield of an MBS sold in the TBA market ($v_k \leq \bar{v}_1$) is $y_1^{\text{TBA}}(v_k) = \frac{v_k}{v_m - h_d} - 1$, which is positive and increases with h_d when $v_k > P_1^{\text{TBA}} = v_m - h_d$. In practice, 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. Hence, through the lens of our model, one component in the TBA yield computed in practice compensates buyers for delivery risk (like in futures contracts), and does not result from the TBA/SP parallel trading environment. The dependence of SP yields on MBS heterogeneity we focus on, however, is tied to the parallel trading.

Second, the dependence of TBA prices on MBS heterogeneity may also arise from a “composition effect.” An increase in h_d , for example, can simply be caused by the issuance of MBSs that are worse than the previously cheapest MBS. This would lead to both an increase in MBS heterogeneity and a drop in TBA prices in the data. A similar composition effect occurs for the SP market, in that the lower TBA price pushes sellers to switch from the TBA market to the SP market, lowering the *average* value of SP MBSs ([Hypothesis 3.1](#)).¹⁶ Empirically, we cannot control for the composition effect on the TBA price because the cheapest MBS (in a relative sense) cannot be held fixed when MBS heterogeneity changes. In contrast, we can control for the composition effect on SP returns by examining SP MBSs with specific characteristics, i.e., holding the v_k fixed in [Eq. \(10\)](#) (see [Section 4](#) for details).

¹⁵This 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.

¹⁶Because the new cheapest MBS and the new set of SP MBSs have worse prepayment characteristics, this composition effect would imply a positive association of MBS heterogeneity with the TBA yield and the *average* SP yield in the data, when prepayment risk premium is nonzero.

4 Data and Measurement

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

MBS coupon cohorts. Our sample contains 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.¹⁷

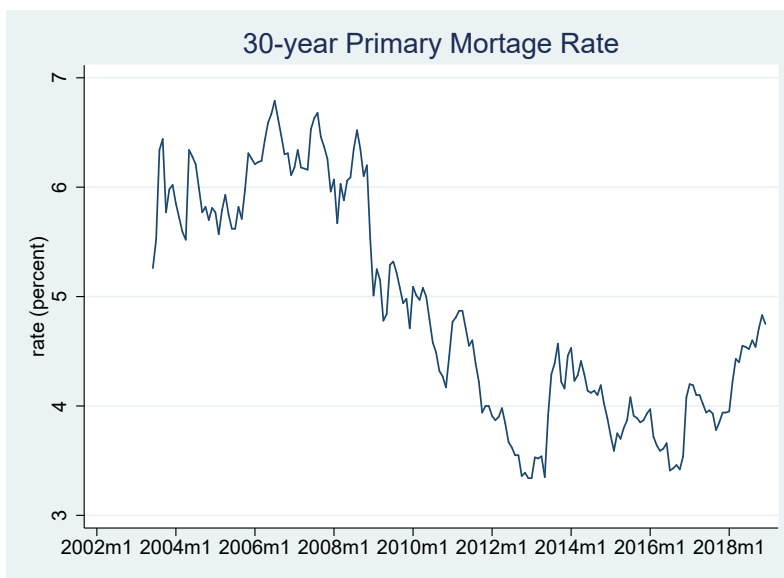
Furthermore, we follow industry practice as described in [Himmelberg, Young, Shan and Henson \(2013\)](#) and also used in [Song and Zhu \(2019\)](#) to exclude the set of MBSs that are least likely to be delivered into TBA contracts. Toward this goal, for each outstanding standard TBA-eligible MBS within each coupon cohort for each month, we obtain the weighted average original FICO score (WAOCS), the weighted average original loan-to-value ratio (WAOLTV), the remaining principal balance (RPB), and the percentage of refinance loans from eMBS through the portal provided Recursion Co.¹⁸ For each coupon cohort for each month, we eliminate MBSs that have at least one of the following characteristics: the refinance share is greater than 75%, the RPB is less than \$150,000, the WAOLTV is above 85%, and the WAOCS is below 680. MBSs with these characteristics “that inhibit efficient prepayments command a price premium, and are not delivered into TBAs” ([Himmelberg et al. \(2013\)](#)).¹⁹ In addition, we exclude cohorts with fewer than 1,000 outstanding MBSs to ensure that we have sufficiently many MBSs to measure cross-sectional heterogeneity.

¹⁷Our choice of moneyness is comparable to that used in other recent studies of MBSs. For example, [Diep et al. \(2019\)](#) focus on MBSs with moneyness in the $[-2.0\%, 3.5\%]$ interval, and [Boyarchenko et al. \(2019\)](#) focus on the $[-2\%, 4\%]$ moneyness interval.

¹⁸These characteristics are calculated based on the “Fixed-Rate Quartile” disclosure files that Fannie Mae began to release in June 2003, using values at issuance and weighting them by the remaining loan balance. For example, the FICO score for each loan underlying an MBS that is used to compute WAOCS is its original value at issuance rather than the value at the time of calculation. But the loan balance used as the weight is the remaining loan balance that is available to investors 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 most often on the fourth business day of each month. Details on the release schedule are available at <http://bit.do/fnm2016>.

¹⁹We did not impose these restrictions on TBA-ineligible securities because, by design, they are not delivered to TBA contracts.

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.

In Panel A of [Table 1](#) 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 the various cohorts. Given the downward trend in mortgage rates in the sample period (as shown in [Figure 2](#)), higher (lower) coupon cohorts appear in the early (late) part of the sample. The time series mean of moneyness increases with the cohort coupon rate, from -0.82% to 2.46% , with the highest and lowest values being 4% and -1.5% , respectively.

In Panel B of [Table 1](#) 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 highest for the 5.5% and 6% cohorts, and is lower 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.

Table 1. Summary Statistics for Monthly CUSIP-Level MBS Characteristics

A: Sample and Moneyiness								
Coupon	Sample			Moneyiness				
	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										
Coupon	WAOCS					SMM				
	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 moneyiness for each coupon cohort. The moneyiness, 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.

Overall, there seem to be enough MBSs within each cohort to measure heterogeneity.

MBS prepayment characteristics. As discussed in [Section 2](#), 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 its monotonic relationship to prepayment risk: a high WAOCS is usually associated with high prepayment risk and low MBS value.²⁰ For each coupon cohort in each month, we use the WAOCS 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). We also obtain the realized prepayment rate for each MBS within each coupon cohort for each month from eMBS, 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.²¹

Panel C of [Table 1](#) 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 $j = 1, \dots, N_{it}$ 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 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 higher coupon cohorts is consistent with a burnout effect

²⁰The weighted average original loan size (WAOSIZE) is another important MBS characteristic that has monotonic relation to prepayment risk. Our main results remain robust using WAOSIZE (see [Section 7](#)).

²¹The SMM can be converted into the annualized constant prepayment rate (CPR) by $CPR = 1 - (1 - SMM)^{12}$.

(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.²²

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% 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 not eligible for TBA trading. Among the TBA-eligible MBSs, higher-LTV groups usually benefit from lower prepayment risk and are of higher value.²³ Using the SPs with fixed characteristics is important because it controls for the composition effect as discussed in Section 3.2.

We match OAS series to the MBS characteristics sample and exclude those without a match. Panel A of Table 2 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. Moreover, within each coupon cohort, the mean OAS is higher for

²²There are several potential issues with OAS measures, which we address in our empirical analyses. First, the dealer's prepayment model can be mis-specified, resulting in errors in OAS measures (Gabaix et al. (2007); Diep et al. (2019)). We conduct robustness checks using the OAS from a different major Wall Street MBS dealer and realized MBS returns. Second, even with a correctly specified prepayment model, the OAS measure contains non-interest-rate prepayment risk premium. We conduct two analyses to show that the effects of MBS heterogeneity are distinct from those of the prepayment risk premium (Section 6). In addition, we obtain OAS series based on the Libor swap curve and for Freddie Mac MBSs and find similar results (Section 7).

²³Consistently, 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 <http://bit.do/sp-summary>.

SPs with higher LTV ratios. This is consistent with our model's implications that low-LTV MBSs benefit more from the existence of TBA trading, enjoying a higher liquidity premium that leads to a lower yield.

In addition, 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.²⁴ Panel C of [Table 2](#) reports the summary statistics the TBA OAS sample. In terms of cohort \times 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 transactions of FNMA 30-year MBSs in the standard coupon cohorts of 2.5%-7%.²⁵ 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 (excluding Mega securities, stripped MBSs, and collateralized mortgage obligations that are backed by existing MBSs). 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.

²⁴In 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.

²⁵Trades 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 2. Summary Statistics of Monthly OAS Series

A: SP Sample						
Coupon	Begin	End	N	Average Moneyiness	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	50.93
3.5	29.50	34.42	36.17	39.76	41.87	54.72
4	38.96	40.67	41.52	41.93	42.48	57.17
4.5	41.26	36.52	40.25	37.12	48.78	63.91
5	27.08	27.32	26.61	26.24	26.59	64.12
C: TBA Sample						
Coupon	Begin	End	N	Average Moneyiness	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 moneyiness 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.

Table 3. Summary of Monthly TBA and SP Trading Activity

A: Sample							
Coupon	Begin	End	N	Average Moneyiness	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						
Coupon	Dollar Volume (\$billion)			Number of Trades		
	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						
Coupon	Dollar Volume (\$billion)			Number of Trades		
	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 moneyiness, 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 \$billion 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.

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 3](#), 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. 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, 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 3](#) 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 and TBA trading is 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 obtain the VIX series from the Chicago Board Options Exchange (CBOE), the balance-sheet-based leverage ratio measure of the aggregate intermediary sector proposed by [He et al. \(2017\)](#), and 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 \(1994\)](#).²⁶ All three variables have been used in the literature to capture the severity of the investment capital constraint, so we use them to proxy for the probability of liquidity shocks in our model (ρ).

All three 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 and “noise” as a parsimonious measure of financial intermediary constraints. 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.

5 Economic Effects of MBS Heterogeneity

In this section, we conduct empirical tests on the economic effects of MBS heterogeneity. We first introduce the measures of MBS heterogeneity and examine their features. We then analyze how MBS heterogeneity affects prices and trading activities of the TBA and SP MBSs based on the testing hypotheses developed in [Section 3.2](#). To be clear, we focus our empirical analysis on variations in MBS heterogeneity *across* coupon cohorts.

5.1 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 3.2](#))—using the prepayment characteristic of the cheapest MBS relative to the average characteristic of all MBSs within a coupon cohort. In particular, we define

$$h_{it}^{\text{WAOCS}} = \text{WAOCS}_{it}^{95\%} - \text{WAOCS}_{it}^{50\%}, \quad (12)$$

²⁶The [Svensson \(1994\)](#) model is used to construct Treasury yield curves that are regular inputs in the Federal Reserve’s policy discussions and publications ([Gurkaynak, Sack and Wright \(2007\)](#)), 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 \(2018\)](#)).

where $WAOCS_{it}^{95\%}$ and $WAOCS_{it}^{50\%}$ 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.²⁷

In Panel A of Table 4 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 1 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 4, 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.

The h_{it}^{WAOCS} measure is based on MBS characteristics, so has the appealing feature of capturing the MBS heterogeneity ex-ante. To show it captures the MBS heterogeneity well ex-post, we construct a measure of heterogeneity using realized prepayment rates directly h_{it}^{SMM} for coupon cohort i at month t as follows:

$$h_{it}^{SMM} = (WAOCS_{it}^{95\%} - WAOCS_{it}^{50\%}) \times ITM + (WAOCS_{it}^{50\%} - WAOCS_{it}^{5\%}) \times OTM.$$

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 By construction, the dispersion measure has a theoretical upper bound of 100% because the highest and lowest possible prepayment rates are 100% and 0%, respectively. In Panel B of Table 4 we report the time-series summary statistics for h_{it}^{SMM} for each coupon cohort i included in our sample. We observe that the average value of h_{it}^{SMM} increases from low to high coupon cohorts.

To verify that the ex-ante measure h^{WAOCS} captures the heterogeneity of future prepayment

²⁷We later conduct robustness checks using (cross-sectional) standard deviation that captures both the downside and upside dispersion. The results are similar.

²⁸A 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 4. 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
C: Regression						
	h^{WAOCS}_{t+1}	$h^{\text{WAOCS}}_{t+1,t+3}$	$h^{\text{WAOCS}}_{t+1,t+12}$	h^{SMM}_{t+1}	$h^{\text{SMM}}_{t+1,t+3}$	$h^{\text{SMM}}_{t+1,t+12}$
h_t^{WAOCS}	0.99*** (198.18)	0.65*** (76.52)	0.54*** (41.97)	1.74*** (3.22)	1.71*** (3.09)	0.78*** (2.92)
Intercept	0.15* (1.71)	0.22 (1.52)	0.30 (1.38)	-29.05*** (-3.16)	-28.62*** (-3.04)	-13.03*** (-2.86)
Obs	1,521	1,461	1,389	1,521	1,497	1,389
R^2_{adj}	1.00	0.99	0.99	0.41	0.42	0.38
Coupon FE	Yes	Yes	Yes	Yes	Yes	Yes

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. Panel C 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 coupon 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 coupon dimensions in parentheses. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.1$, where p is the p-value.

rates well, we consider the following regression:

$$h_{i,t+n}^{\text{SMM}} = \sum_i \alpha_t D_i + \beta \cdot h_{it}^{\text{WAOCS}} + \varepsilon_{it}, \quad (13)$$

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 cohort fixed effect D_i 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.

In the first three columns of Panel C of [Table 4](#), we report results of the regression ([Eq. \(13\)](#)) 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 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 . The adjusted R^2 is much larger than that for forecasting $h_{i,t+n}^{\text{SMM}}$, suggesting that $h_{i,t}^{\text{WAOCS}}$ is highly persistent over time for each cohort i . Overall, the results show that investors can form reasonably accurate expectations of future MBS heterogeneity, which is consistent with the model setup in [Section 3](#).

5.2 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 :

$$\text{OAS}_{itj} = \sum_t \alpha_t D_t + \beta_1 \cdot h_{it}^{\text{WAOCS}} + \beta_2 \cdot \text{SMM}_{itj} + \varepsilon_{itj}, \quad (14)$$

for each j , where j represents one of the six types of SPs based on LTV ratios. We control for the prepayment rate SMM_{itj} . Time fixed-effects are included, so the coefficient β_1 captures the effects of MBS heterogeneity on the cross-sectional variation of OAS. In Panel A of [Table 5](#) we report the results of the panel regression ([Eq. \(14\)](#)) for SP MBSs with LTV ratios higher than

105% in the first two columns and for SP MBSs with LTV ratios lower than 105% in the last four columns. The t-statistics using standard errors two-way clustered along the time and coupon dimensions are reported in parentheses.

We observe that h^{WAOCS} significantly affects SP OAS positively, consistent with our model's prediction that having future TBA trading as an option affects current SP prices. We further observe that the effects of h^{WAOCS} on the OAS are weaker for TBA-eligible SP MBSs with higher LTV ratios that are less likely to be delivered into TBA contracts than for those with lower LTV ratios that are more likely to be delivered into TBA contracts, and even weaker for SP MBSs with LTV ratios higher than 105% that are ineligible for TBA delivery. This pattern squares with our model's prediction that the effects of MBS heterogeneity on SP MBSs' returns are weaker for MBSs that are less likely to be traded in the TBA market in the future.

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 16 ($\approx 11.81 \times 1.33$) basis points for SP MBSs with LTV ratios in the 80-90% range, and by about 10 ($\approx 11.81 \times 0.87$) basis points for SP MBSs with LTV ratios in the 100%-105% range. That is, the effects diminish by almost half for SP MBSs that are unlikely to be delivered into TBA contracts. In addition, the effect is about 8 ($\approx 11.81 \times 0.67$) basis points for TBA-ineligible SP MBSs with LTV ratios in the 105%-125% range, much like those with LTV ratios in the 100%-105% range. This is possibly because these MBSs could become eligible for TBA delivery once high LTV-ratio loans default and drop out of the pool.

To check the overall explanatory power of MBS heterogeneity, we reports panel regression without h^{WAOCS} in Panel B of [Table 5](#). We observe that the increase in adjusted R^2 when including h^{WAOCS} ranges from about 6% to 15%. 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 \(1986\)](#)). As [Gao et al. \(2017\)](#) show, the SP trading cost does decrease with the LTV ratio, but it increases substantially across the 105% threshold. The weaker effects of h^{WAOCS} on SP MBSs with LTV ratios higher than 105% seem inconsistent with this alternative interpretation.

In addition, although our main focus is the effect of MBS heterogeneity on SP yields, we also run the regression ([Eq. \(14\)](#)) for TBA MBSs. We find that h^{WAOCS} significantly affects OAS of TBA MBSs positively, consistent with the generic CTD discount. A one-standard-deviation increase

Table 5. MBS Heterogeneity and Yields

	TBA Ineligible SP (LTV)		TBA-Eligible SP (LTV)			
	105-125	> 125	80-90	90-95	95-100	100-105
A: Regression on h^{WAOCs} and SMM						
h^{WAOCs}	0.67*	0.83**	1.33***	1.27***	1.06**	0.87*
	(1.67)	(2.26)	(4.24)	(3.30)	(2.45)	(1.95)
SMM	-2.76***	-1.21**	-1.50***	-1.88***	-2.01***	-2.18***
	(-3.24)	(-2.18)	(-4.14)	(-5.03)	(-4.37)	(-4.74)
Intercept	62.25***	76.88***	14.63***	29.20***	38.07***	45.92***
	(6.49)	(8.15)	(4.75)	(5.50)	(6.05)	(6.82)
Obs	390	390	390	390	390	390
R^2_{adj}	0.52	0.53	0.64	0.63	0.56	0.54
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
B: Regression on SMM						
SMM	-1.95**	0.01	-0.65**	-1.00***	-1.24***	-1.51***
	(-2.40)	(0.02)	(-2.48)	(-3.80)	(-3.87)	(-4.65)
Intercept	70.97***	87.74***	29.50***	44.88***	51.94***	57.83***
	(10.13)	(15.21)	(19.42)	(13.91)	(14.30)	(15.91)
Obs	390	390	390	390	390	390
R^2_{adj}	0.49	0.43	0.47	0.47	0.45	0.48
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: In this table we report the results for panel regressions of the OASs of two groups of TBA-ineligible SP MBSs (first two columns) and four groups of TBA-eligible SP MBSs (last four columns), on h^{WAOCs} for FNMA 30-year MBSs based on monthly data. In Panel A we report the results for bi-variate regressions controlling for SMM, while Panel B reports uni-variate regressions on SMM only. Time dummies are included, and t-statistics based on robust standard errors that are two-way clustered along the time and coupon 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.

of h^{WAOCs} across coupon cohorts is associated with an increase in OAS by about 19 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.

5.3 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 4, we use the VIX and Distress measures to proxy for the probability of liquidity shocks (ρ in our model), as these variables are shown to capture the extent of investment capital constraints. We consider the following panel regression:

$$\text{OAS}_{itj} = \sum_t \alpha_t D_t + \sum_j \alpha_j D_j + \beta_1 \cdot h_{it}^{\text{WAOCs}} + \beta_2 \cdot h_{it}^{\text{WAOCs}} \times \rho_t + \beta_3 \cdot \rho_t + \gamma \cdot \text{SMM}_{itj} + \varepsilon_{itj}, \quad (15)$$

for the whole SP sample by pooling all six types of SP MBSs together. We control for the prepayment rate SMM_{itj} 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 report in columns (1)-(2) of Table 6 regressions with VIX and Distress proxying for ρ_t , respectively. The coefficient on the interaction term $h_{it}^{\text{WAOCs}} \times \rho_t$ is positive and highly significant, confirming that the effects of MBS heterogeneity on SP yields are stronger when selling pressure is higher. In columns (3)-(4) we report similar regressions with time fixed-effects to absorb all time-series variables. The interaction term using VIX has marginal statistical significance, which is not surprising because VIX is expected to be only marginally relevant to MBS markets. In contrast, the interaction term using Distress is still highly significant, likely because it is based on balance-sheet data of large financial intermediaries closely related to MBS markets.

5.4 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).

Table 6. Liquidity Shocks

	(1)	(2)	(3)	(4)
h^{WAOCS}	0.25 (0.61)	1.51*** (3.55)	0.74 (1.25)	1.44*** (2.69)
SMM	-0.63* (-1.85)	-0.81*** (-2.60)	-0.60 (-1.44)	-0.73 (-1.55)
$h^{\text{WAOCS}} \times \text{VIX}$	0.07*** (5.24)		0.04* (1.66)	
VIX	-0.82* (-1.83)			
$h^{\text{WAOCS}} \times \text{Distress}$		0.71*** (9.03)		0.65*** (4.17)
Distress		-8.83*** (-2.93)		
Intercept	18.04 (1.60)	6.35 (0.52)	89.74 (0.40)	11.40 (0.28)
Observations	7,696	5,752	7,696	5,752
R^2_{adj}	0.35	0.42	0.53	0.57
Time FE	No	No	Yes	Yes
SP Type FE	Yes	Yes	Yes	Yes

Note: We report in this table the results for panel regressions of the SP OASs on the interaction terms $h^{\text{WAOCS}} \times \text{VIX}$ and $h^{\text{WAOCS}} \times \text{Distress}$ 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 regressions reported in columns (1) - (2) where VIX and Distress are controlled for directly, while are included in the regressions reported in columns (3) - (4) to control for time-series variations. All regressions include SMM as a control. The t-statistics based on robust standard errors that are two-way clustered along the time and coupon 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.

In Columns (1)-(2) of Panels A of [Table 7](#), 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 MBSs²⁹. 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 increases. Further, the regression of the 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 the MBS heterogeneity increases, the fraction of MBSs sold through SP rather than TBA markets increases because greater TBA price discounts prompt sellers of high-quality 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 the magnitudes of the regression coefficients are appreciably lower for dealer-customer trades than for all trades of on the TBA market (but not on the SP market). 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 included 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 SP/TBA ratio of trading volume.

6 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

²⁹The results controlling for outstanding balance are similar, as presented in [Section 7](#).

Table 7. 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*** (-7.58)	-0.30*** (-2.93)	0.11*** (7.67)	-234.71*** (-7.10)	-62.15*** (-3.32)	0.10*** (8.77)
Issuance	10.82*** (7.01)	0.66*** (6.22)	0.02 (1.57)	477.51*** (8.47)	4.58 (0.21)	-0.03*** (-3.32)
Intercept	290.94*** (7.03)	14.30*** (3.51)	-5.97*** (-16.33)	11,519.83*** (6.63)	3,379.04*** (3.56)	-3.88*** (-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*** (-9.11)	-0.25*** (-3.12)	0.10*** (7.74)	-43.32*** (-6.46)	-44.60*** (-3.54)	0.07*** (8.41)
Issuance	5.19*** (7.66)	0.64*** (7.66)	0.01* (1.67)	116.50*** (9.20)	8.17 (0.58)	-0.04*** (-6.36)
Intercept	105.79*** (9.15)	11.36*** (3.69)	-4.88*** (-16.31)	1,761.97*** (6.23)	2,247.09*** (3.67)	-1.65*** (-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 ratios on h^{WAOCs} for FNMA 30-year MBS 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 coupon 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.

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.

6.1 Premium and Discount Markets

As analyzed in [Gabaix et al. \(2007\)](#) and [Diep et al. \(2019\)](#), 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, the effects of MBS heterogeneity arise from the unique market structure of MBS trading, as illustrated in our model. MBS heterogeneity always affects MBS yields positively, whether the MBS market is in premium or in discount. 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 8](#) 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.³⁰ 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 moneyness, as shown in [Diep et al. \(2019\)](#), this result implies that the relationship between MBS heterogeneity and prepayment risk exposure is unlikely to change signs

³⁰To measure market type, we follow the method of [Diep et al. \(2019\)](#). 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.

across premium and discount markets.

Then we report panel regressions of the OAS on h^{WAOCs} in Panel B of [Table 8](#), 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 5.3](#). The regression coefficients on h^{WAOCs} are significantly positive regardless of market type. Compared with regressions with SMM only, the incremental R^2 of h^{WAOCs} is about 5%. Overall, these results show that the effects of MBS heterogeneity are distinct from the effects of exposure to market-level prepayment risk.

6.2 IO and PO Strips

Many studies have focused on studying the effects of individual-security-level prepayment risks. For example, a recent comprehensive analysis conducted by [Boyarchenko et al. \(2019\)](#) using IO and PO strips shows that the non-interest-rate prepayment risk premium has significant explanatory power for MBS yields across coupon cohorts. An intriguing feature of IO and PO strips is that their cash flows have opposite exposure to the same prepayment risk (of the same underlying collateral). Intuitively, this is because prepayments reduce total interest payments and accelerate principal payments. [Boyarchenko et al. \(2019\)](#) use this feature to pin down the prepayment risk premium component in the OAS. Instead, we use this feature to differentiate the effects of MBS heterogeneity from that of the individual-security-level prepayment risk. We shall show that, unlike the prepayment risk, MBS heterogeneity raises returns on both IO and PO strips.

To achieve this goal, 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).³¹ 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 1](#)) at the cohort-month level. We also obtain characteristics of the collateral securities from eMBS.

³¹As of June 3, 2019, all TBA-eligible Megas, regardless of issue date, are labeled as “Major Supers”. Details are provided at <http://bit.do/fnm-megas>.

Table 8. 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)		(103.36)	
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}	1.50***		1.53***		0.76***	
	(5.03)		(4.77)		(3.19)	
SMM	0.47	-0.66*	0.46	-0.68*	1.99***	0.91
	(1.54)	(-1.90)	(1.44)	(-1.94)	(3.07)	(0.92)
Intercept	28.27***	-14.10	27.60***	-15.54	27.76***	13.23**
	(9.77)	(-1.16)	(8.61)	(-1.21)	(12.06)	(2.28)
Obs	7,696	7,696	7,111	7,111	585	585
R^2_{adj}	0.46	0.52	0.43	0.49	0.85	0.86
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
SP Type EF	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 coupon 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.

Table 9. Summary of IO/PO Trusts

Trust number	Collateral coupon	N	Sample		Characteristics at Issuance				Time Series Means		
			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.

In [Table 9](#) we provide summary statistics for IO/PO trusts. 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 the other 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 1](#). 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%.

To study how MBS heterogeneity is associated with the OAS of IO/PO strips, we construct respective 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 of [Table 10](#), 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 cohort. 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 10](#), 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 of IO/PO strips. This would happen when the value of the underlying collateral is not among the highest; otherwise, they would always be sold on the SP market and their prices do 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. We then match them

Table 10. 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*** (4.20)		5.34*** (6.31)	11.06*** (6.43)		12.83*** (6.98)	
SMM		-2.64 (-0.72)	-27.40*** (-3.91)		23.89** (2.23)	-35.54* (-1.89)	
Intercept	-207.79*** (-5.22)	-50.29*** (-3.55)	-167.34*** (-6.85)	-472.14*** (-6.16)	-138.66*** (-4.96)	-419.65*** (-6.64)	
Obs	612	612	612	612	612	612	
R^2_{adj}	0.79	0.77	0.79	0.86	0.82	0.86	
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	
D: MBS 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.

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 10](#) 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.

Overall, our results complement [Gabaix et al. \(2007\)](#), [Boyarchenko et al. \(2019\)](#), [Chernov et al. \(2017\)](#) and [Diep et al. \(2019\)](#) in delivering an in-depth picture of MBS pricing: both fundamental prepayment risk and market structure of trading are essential economic forces driving the cross-sectional variations in MBS returns.

7 Robustness

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

First, the data sample in the main analysis of [Section 5](#) excludes cohorts with fewer than 1,000 MBSs. In Panel A of [Table 11](#), we report all the main results (those reported in [Table 4](#), [Table 5](#), [Table 6](#), and [Table 7](#)) 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 11](#) 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 two alternative measures of MBS heterogeneity. The first uses the difference between the 90th percentile and the median of WAOCS, denoted as $h^{\text{WAOCS},10\%}$. The second is similar to the baseline measure, using the difference between the 95th percentile and the median, but of WAOSIZE, which is another important MBS characteristic, denoted as $h^{\text{WAOSIZE},5\%}$. In Panels A and B of [Table 12](#) we report results from the main analysis using $h^{\text{WAOCS},10\%}$ and $h^{\text{WAOSIZE},5\%}$, respectively. Our main results are robust using these two alternative measures of MBS heterogeneity.

Fourth, a few studies, such as [Fabozzi and Mann \(2011\)](#) and [Belikoff, Levin, Stein and Tian](#)

Table 1.1. 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}	105-125	> 125	80-90	90-95	95-100	100-105	SP	Volume	# Trade	
h_t^{WAOCS}	1.74** (3.13)	0.99*** (193.44)	0.67* (1.67)	0.83** (2.26)	1.33*** (4.24)	1.27*** (3.30)	1.06** (2.45)	0.87* (1.95)	0.25 (0.61)	0.11*** (7.75)	0.10*** (8.28)	
SMM			-2.76*** (-3.24)	-1.21** (-2.18)	-1.50*** (-4.14)	-1.88*** (-5.03)	-2.01*** (-4.37)	-2.18*** (-4.74)	-0.63* (-1.85)			
$h_t^{WAOCS} \times VIX$								0.07*** (5.25)				
VIX									-0.81* (-1.82)			
Issuance										0.02 (1.64)	-0.03*** (-3.13)	
Intercept	-13.92 (-0.99)	0.20* (1.95)	62.25*** (6.49)	76.88*** (8.15)	14.63*** (4.75)	29.20*** (5.50)	38.07*** (6.05)	45.92*** (6.82)	18.00 (1.60)			
Obs	1,411	1,411	390	390	390	390	390	390	7,684	368	368	
R_{adj}^2	0.40	1.00	0.52	0.53	0.64	0.63	0.56	0.54	0.34	0.69	0.81	
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	

B: FHLMC 30-year MBS

	One-month Ahead		SP OAS						Liquidity Shock		SP/TBA Ratio	
	h_{t+1}^{SMM}	$h_{t+1}^{WAOCSIZE}$	105-125	> 125	80-90	90-95	95-100	100-105	OAS	Volume	# Trade	
$h_t^{WAOCSIZE}$	1.84*** (3.62)	0.99*** (254.07)	0.66 (1.48)	0.96** (2.18)	1.21*** (3.53)	1.19*** (2.78)	0.99** (2.10)	0.81 (1.65)	0.71 (1.18)	0.09*** (5.41)	0.05*** (3.50)	
SMM			-2.38*** (-2.77)	-1.00 (-1.55)	-1.26*** (-3.42)	-1.65*** (-4.21)	-1.74*** (-3.60)	-1.92*** (-3.92)	-0.18 (-0.66)			
$h_t^{WAOCSIZE} \times VIX$									0.03* (1.73)			
VIX									-0.00 (-0.00)			
Issuance										-0.04*** (-3.22)	-0.05*** (-2.74)	
Intercept	-20.04* (-1.65)	0.15 (1.49)	54.40*** (5.68)	62.87*** (89.86)	22.93*** (2.90)	30.79*** (3.34)	38.47*** (3.51)	42.70*** (3.37)	9.66 (0.62)			
Obs	1,443	1,443	388	388	388	388	388	388	6,690	360	360	
R_{adj}^2	0.45	1.00	0.49	0.52	0.55	0.54	0.49	0.49	0.21	0.55	0.57	
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	

Note: In the first two columns we report the results for panel regressions of one-month-ahead heterogeneity measures (h_{t+1}^{SMM} and $h_{t+1}^{WAOCSIZE}$) 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 $h_{it}^{WAOCS} \times VIX$ 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 MBSSs, 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 $h_{it}^{WAOCS} \times VIX$ that, however, include fixed effects for SP type. The t-statistics based on robust standard errors that are two-way 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 12. Alternative Measures of MBS Heterogeneity

A: 10th percentile													
One-month Ahead		SP OAS						Liquidity Shock		SP/TBA Ratio			
h_{t+1}^{SMM}	h_{t+1}^{WAOCS}	105-125	> 125	80-90	90-95	95-100	100-105	SP	Volume	# Trade			
h_t^{WAOCS}	2.20*** (3.26)	0.90 (1.62)	1.05** (2.21)	1.55*** (3.89)	1.49*** (3.04)	1.25** (2.28)	1.05* (1.83)	0.17 (0.33)	0.14*** (7.16)	0.14*** (8.13)			
SMM		-2.77*** (-3.14)	-1.14** (-2.01)	-1.38*** (-3.87)	-1.75*** (-4.84)	-1.90*** (-4.25)	-2.09*** (-4.68)	-0.56 (-1.51)					
$h_t^{WAOCS} \times VIX$								0.10*** (5.04)					
VIX								-0.97* (-1.82)					
Issuance									0.02 (1.49)	-0.03*** (-3.34)			
Intercept	-29.72*** (-3.25)	62.31*** (6.86)	77.65*** (9.08)	16.51*** (8.44)	31.11*** (6.62)	39.66*** (7.00)	47.10*** (7.65)	20.86* (1.82)	-6.00*** (-15.25)	-3.92*** (-12.14)			
Obs	1,523	390	390	390	390	390	390	7,696	377	377			
R_{adj}^2	0.42	0.52	0.53	0.63	0.61	0.55	0.54	0.34	0.67	0.81			
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes			

B: WAOSIZE													
One-month Ahead		SP OAS						Liquidity Shock		SP/TBA Ratio			
h_{t+1}^{SMM}	h_{t+1}^{WAOCS}	105-125	> 125	80-90	90-95	95-100	100-105	OAS	Volume	# Trade			
h_t^{WAOCS}	25.08** (2.26)	8.98 (1.32)	11.30*** (2.68)	21.81*** (3.76)	16.17*** (2.93)	12.34** (1.97)	8.75 (1.58)	-19.37 (-1.23)	3.71*** (2.75)	4.12*** (3.84)			
SMM		-2.00** (-2.48)	-0.02 (-0.03)	-0.69*** (-2.77)	-1.04*** (-4.10)	-1.26*** (-3.88)	-1.52*** (-4.65)	0.01 (0.01)					
$h_t^{WAOCS} \times VIX$								3.35** (2.48)					
VIX								-1.87 (-1.24)					
Issuance									-0.05*** (-5.38)	-0.09*** (-11.51)			
Intercept	-18.67** (-2.12)	15.41** (2.23)	7.84* (1.67)	39.53*** (4.98)	49.02*** (6.45)	62.05*** (6.32)	76.39*** (8.67)	36.18** (2.12)	-6.56*** (-4.74)	-5.12*** (-4.43)			
Obs	1,521	1,445	390	390	390	390	390	7,696	377	377			
R_{adj}^2	0.38	0.67	0.55	0.48	0.49	0.50	0.46	0.26	0.44	0.61			
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes			

Note: In the first two columns we report results for panel regressions using monthly data of one-month-ahead heterogeneity measures (h_{t+1}^{SMM} and h_{t+1}^{WAOCS}) on h_t^{WAOCS} . In the next fix columns we 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 $h_t^{WAOCS} \times VIX$ 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_t^{WAOCS} , while controlling for monthly total new issuance (\$billion). For Panel A we use the h_{t+1}^{WAOCS} measure computed as the difference between the 90th percentile and the median of WAOCS, while for Panel B we use the h_{t+1}^{WAOCS} measure computed as the difference between the 95th percentile and the median of WAOCS. Time fixed-effects are included in all but the regressions on the interaction term $h_t^{WAOCS} \times VIX$, however, include fixed-effects for SP type. 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.

(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 13](#) we repeat the main analyses (those reported in [Table 5](#) and [Table 6](#)) 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. \(2019\)](#). Regression results for these alternative OAS and return series are reported in Panel B of [Table 13](#). 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 14](#) reports regression results of MBS trading activities on h^{WAOCS} , by controlling for outstanding balance instead of new issuance. The results are similar.

8 Conclusion

To the best of our knowledge, this paper conducts the first analysis of the distinctive asset pricing effects of the unique TBA/SP parallel trading environment. We construct a simple “liquidity-based asset pricing” model that allows investors to choose between TBA and SP trading. Two distinctive economic channels arise from the interaction between TBA and SP markets: first, high-value MBSs are more likely to be sold in the SP market; second, buyers in the SP market can use TBA as a backup selling venue in the future when they experience liquidity shocks.

Measuring the dispersion of MBS values within a cohort based on individual-MBS-level prepayment characteristics, denoted as MBS heterogeneity, we empirically confirm three main sets of implications. First, although CTD issues are absent from SP trading by definition, MBS heterogeneity associated with CTD discounts affects SP returns positively, with the effect stronger for lower-value SPs. Second, high selling pressure amplifies the effects of MBS heterogeneity on SP returns. Third, greater MBS heterogeneity dampens SP and TBA trading activities but

Table 13. Alternative Measures of MBS Yields

	A: Libor OAS						
	105-125	> 125	80-90	90-95	95-100	100-105	ALL SP
h_t^{WAOCS}	1.10*** (3.38)	1.28*** (4.52)	1.47*** (4.96)	1.43*** (4.55)	1.28*** (3.72)	1.13*** (3.37)	0.30 (0.85)
SMM	-2.47*** (-2.78)	-1.14* (-1.81)	-1.76*** (-4.76)	-2.00*** (-6.18)	-2.02*** (-5.10)	-2.03*** (-5.69)	-0.13 (-0.51)
$h_t^{\text{WAOCS}} \times \text{VIX}$							0.04*** (2.82)
VIX							-0.43 (-0.91)
Intercept	52.52*** (7.92)	67.65*** (9.09)	13.06*** (5.17)	25.88*** (5.15)	32.90*** (5.81)	39.48*** (6.79)	8.88 (0.97)
Obs	384	384	384	384	384	384	7,559
R^2_{adj}	0.62	0.75	0.68	0.68	0.66	0.69	0.19
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	No
B: Dealer OAS and Hedged Returns							
	OAS		Hedged Returns				
h_t^{WAOCS}	1.12*** (3.61)	0.70 (1.39)			0.061*** (4.809)	-0.004 (-0.219)	
SMM	0.14 (0.74)	-0.29 (-1.15)			-0.012 (-1.183)	0.021*** (3.347)	
$h_t^{\text{WAOCS}} \times \text{VIX}$		0.04** (2.19)				0.001* (1.737)	
VIX		0.05 (0.08)				-0.071** (-2.303)	
Intercept	-15.51** (-2.47)	8.58 (0.56)			-0.344 (-0.712)	0.959 (1.441)	
Obs	1,107	1,107			1,100	1,100	
R^2_{adj}	0.81	0.27			0.746	0.081	
Time FE	Yes	No			Yes	No	

Note: In the first six columns of Panel A we report the results for panel regressions of respective OASs of six group of FNMA 30-year SP MBSs on h^{WAOCS} , while in the last column we report the panel regression on the interaction term $h^{\text{WAOCS}} \times \text{VIX}$ 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 $h^{\text{WAOCS}} \times \text{VIX}$ separately. All regressions control for SMM. The t-statistics based on robust standard errors that are two-way clustered along the time and coupon 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 14. 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*** (-3.810)	-0.483*** (-3.112)	0.060*** (6.229)	-352.524*** (-3.959)	-15.891 (-1.216)	0.099*** (12.417)
Outstanding	0.152 (1.284)	0.007 (0.865)	-0.005*** (-4.995)	7.041 (1.450)	6.132*** (9.956)	-0.002* (-1.806)
Intercept	360.217*** (3.619)	19.495** (2.563)	-3.289*** (-12.367)	14,429.979*** (3.625)	616.343 (0.949)	-3.533*** (-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*** (-3.906)	-0.453*** (-3.258)	0.052*** (5.289)	-67.384*** (-3.813)	-18.587* (-1.882)	0.080*** (10.299)
Outstanding	0.081* (1.699)	0.004 (0.577)	-0.005*** (-4.425)	2.317** (2.514)	3.707*** (7.606)	-0.001 (-0.996)
Intercept	135.150*** (3.537)	17.673*** (2.606)	-2.468*** (-9.591)	2,196.159*** (3.055)	646.539 (1.410)	-1.834*** (-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 coupon 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.

increases their ratio. 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.

Appendices

A Additional Equilibrium Types

In this appendix, we discuss two additional types of equilibrium that complement the equilibrium type analyzed in [Proposition 2](#) in making a complete equilibrium analysis.

Proposition 3 (Time-1 Equilibrium). *If $C_1^{\text{SP}} < (1 - \rho)c_{2,\ell}$, then $\bar{v}_1 = v_m - h_d + \frac{C_1^{\text{SP}}}{1 - \rho} < \bar{v}_{2,\ell}$. At time 1, if the value of an MBS $v_k \geq \bar{v}_1$, its seller sells the MBS in the SP market at price*

$$P_1^{\text{SP}} = \begin{cases} v_k - \rho \mathbf{E}[C_2^{\text{SP}}] & \text{if } v_k > \bar{v}_{2,h}, \\ (1 - \rho\pi_h)v_k + \rho\pi_h(v_m - h_d) - \rho(1 - \pi_h)c_{2,\ell} & \text{if } \bar{v}_{2,\ell} \leq v_k \leq \bar{v}_{2,h}, \\ (1 - \rho)v_k + \rho(v_m - h_d) & \text{if } \bar{v}_1 \leq v_k \leq \bar{v}_{2,\ell} \end{cases} \quad (\text{A1})$$

At time 1, if the value of an MBS $v_k < \bar{v}_1$, it is sold in the TBA market at price $v_m - h_d$.

Proposition 4 (Time-1 Equilibrium). *If $C_1^{\text{SP}} > c_{2,h} - \rho \mathbf{E}[C_2^{\text{SP}}]$, then $\bar{v}_1 = C_1^{\text{SP}} + \rho \mathbf{E}[C_2^{\text{SP}}] > \bar{v}_{2,h}$ and any MBS sold on the SP market at time 1 would never be sold on the TBA market at time 2. At time 1, if the value of an MBS $v_k \geq \bar{v}_1$, its seller sells the MBS on the SP market at price*

$$P_1^{\text{SP}} = v_k - \rho \mathbf{E}[C_2^{\text{SP}}] \quad (\text{A2})$$

At time 1, if the value of an MBS $v_k < \bar{v}_1$, it is sold on the TBA market at price p_1^{TBA} .

Most generally, we can write the TBA threshold at time 1 as follows:

$$\bar{v}_1 = v_m - h_d + \min \left\{ C_1^{\text{SP}} + \rho \mathbf{E}[C_2^{\text{SP}}], \frac{C_1^{\text{SP}} + \rho(1 - \pi_h)c_{2,\ell}}{1 - \rho\pi_h}, \frac{C_1^{\text{SP}}}{1 - \rho} \right\}. \quad (\text{A3})$$

[Figure 3](#) illustrates this result. We can see that, depending on the level of C_1^{SP} , the realized equilibrium is one of the equilibria described in [Proposition 2](#), [Proposition 3](#), or [Proposition 4](#).

[Figure 3](#) also demonstrates the benefits of allowing TBA trading at time 2. If there is no TBA market at time 2, then, at time 1, $\bar{v}_1 = v_m - h_d + C_1^{\text{SP}} + \rho \mathbf{E}[C_2^{\text{SP}}]$, which is plotted with the orange dashed line. The existence of the time-2 TBA market *reduces* coverage of the time-1 TBA market because it enables some SP sellers to avoid paying the high SP costs at time 2 and increases time-1 SP buyers' willingness-to-pay. Hence, a time-1 seller can obtain higher revenue in the SP market and is less likely to choose the TBA

market. The difference between the blue and orange lines in Figure 3 reflects the benefit of the time-2 TBA market for time-1 sellers.

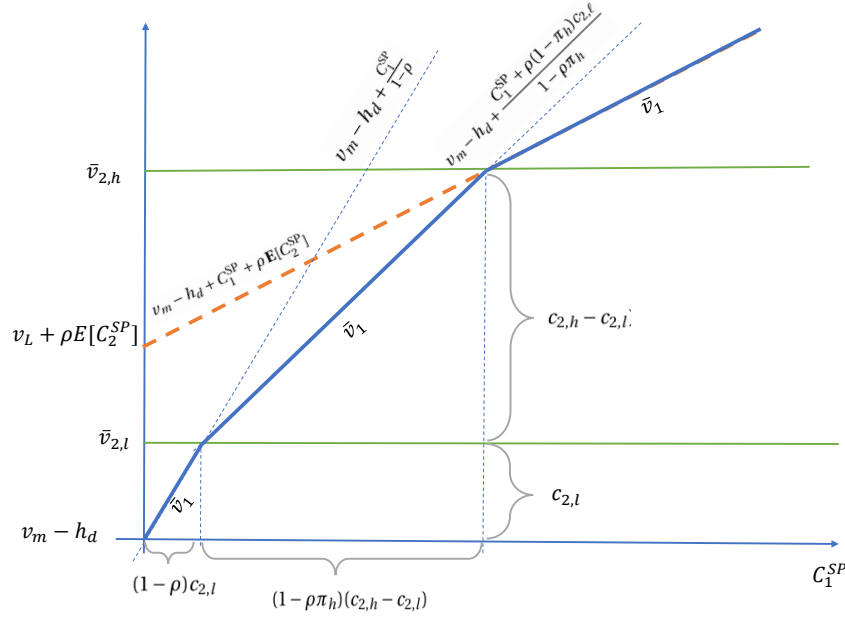


Figure 3. Upper bound of time-1 TBA MBS value \bar{v}_1

B Proofs

Proof of Proposition 1. Buyers in the SP market know that they can sell an asset at time 3 at price v . Because the interest rate is zero, $P_2^{\text{SP}}(v) = v$. Thus, a seller may obtain revenue of $v_k - C_2^{\text{SP}}$ in the SP market or $P_2^{\text{TBA}} = v_m - h_d$ in the TBA market. Hence, she chooses the TBA market if and only if $v_m - h_d \geq v_k - C_2^{\text{SP}}$, which is equivalent to $v_k \leq v_m - h_d + C_2^{\text{SP}} = \bar{v}_2$. \square

Proof of Lemma 1. At time 2, the seller of an MBS of value v_k can obtain revenue of

$$\max\{v_k - C_2^{\text{SP}}, v_m - h_d\} = \begin{cases} v_k - C_2^{\text{SP}} & \text{if } v_k > \bar{v}_{2,h} \\ v_k - c_{2,\ell} & \text{if } v_k \in [\bar{v}_{2,\ell}, \bar{v}_{2,h}] \text{ and } C_2^{\text{SP}} = c_{2,\ell} \\ v_m - h_d & \text{if } v_k \in [\bar{v}_{2,\ell}, \bar{v}_{2,h}] \text{ and } C_2^{\text{SP}} = c_{2,h} \text{ or if } v_k < \bar{v}_{2,\ell} \end{cases} \quad (\text{B1})$$

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

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

□

Proof of Proposition 2. \bar{v}_1 satisfies $P_1^{\text{SP}}(\bar{v}_1) - C_1^{\text{SP}} = P_1^{\text{TBA}} = v_m - h_d$. Then, by Lemma 1,

$$C_1^{\text{SP}} = P_1^{\text{SP}}(\bar{v}_1) - v_m + h_d = \begin{cases} \bar{v}_1 - v_m + h_d - \rho \mathbf{E}[C_2^{\text{SP}}] & \text{if } \bar{v}_1 > \bar{v}_{2,h} \\ (1 - \rho)\bar{v}_1 + \rho(\pi_h(v_m - h_d) + (1 - \pi_h)(\bar{v}_1 - c_{2,\ell})) - v_m + h_d & \text{if } \bar{v}_{2,\ell} \leq \bar{v}_1 \leq \bar{v}_{2,h} \\ (1 - \rho)(\bar{v}_1 - v_m + h_d) & \text{if } \bar{v}_1 < \bar{v}_{2,\ell} \end{cases} \quad (\text{B2})$$

If $\bar{v}_{2,\ell} \leq \bar{v}_1 \leq \bar{v}_{2,h}$, Eq. (B2) implies that $\bar{v}_1 = v_m - h_d + \frac{C_1^{\text{SP}} + \rho(1 - \pi_h)c_{2,\ell}}{1 - \rho\pi_h}$. It holds if $c_{2,\ell} \leq \frac{C_1^{\text{SP}} + \rho(1 - \pi_h)c_{2,\ell}}{1 - \rho\pi_h} \leq c_{2,h}$, or equivalently $(1 - \rho)c_{2,\ell} \leq C_1^{\text{SP}} \leq c_{2,h} - \rho \mathbf{E}[C_2^{\text{SP}}]$. □

Proof of Proposition 3. If $\bar{v}_1 > \bar{v}_{2,h}$, Eq. (B2) implies that $\bar{v}_1 = v_m - h_d + C_1^{\text{SP}} + \rho \mathbf{E}[C_2^{\text{SP}}]$. It follows that $v_m - h_d + C_1^{\text{SP}} + \rho \mathbf{E}[C_2^{\text{SP}}] > \bar{v}_{2,h} = v_m - h_d + c_{2,h}$, which holds if and only if $C_1^{\text{SP}} > c_{2,h} - \rho \mathbf{E}[C_2^{\text{SP}}] = (1 - \rho\pi_h)(c_{2,h} - c_{2,\ell}) + (1 - \rho)c_{2,\ell}$. □

Proof of Proposition 4. If $\bar{v}_1 < \bar{v}_{2,\ell}$, then $C_1^{\text{SP}} = (1 - \rho)(\bar{v}_1 - v_m + h_d)$, which implies that $\bar{v}_1 = v_m - h_d + \frac{C_1^{\text{SP}}}{1 - \rho}$. It can hold only if $C_1^{\text{SP}} < (1 - \rho)c_{2,\ell}$. □

References

- Akerlof, G.A.**, “The Market for Lemons: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics*, 1970, 84, 488–500.
- Amihud, Yakov and Haim Mendelson**, “Asset pricing and the bid-ask spread,” *Journal of Financial Economics*, 1986, 17 (2), 223 – 249.
- , ———, and **Lasse Heje Pedersen**, “Liquidity and Asset Prices,” *Foundations and Trends in Finance*, 2006, 1 (4), 269–364.
- An, Yu, Wei Li, and Zhaogang Song**, “Cohort Trading and Security Design: Theory and Evidence from Agency MBS Markets,” *Johns Hopkins Carey Business School, working paper*, 2020.
- Bartolini, Leonardo, Spence Hilton, Suresh Sundaresan, and Christopher Tonetti**, “Collateral Values By Asset Class: Evidence from Primary Securities Dealers,” *Review of Financial Studies*, 2011, 24(1), 248–278.
- Belikoff, Alexander, Kirill Levin, Harvey Stein, and Xusheng Tian**, “Analysis of Mortgage-Backed Securities: Before and After the Credit Crisis,” *Bloomberg LP, Version 2.2*, 2010.
- Bessembinder, Hendrik, Chester Spatt, and Kumar Venkataraman**, “A Survey of the Microstructure of Fixed-Income Markets,” *Journal of Financial and Quantitative Analysis*, 1 2019.
- , **William Maxwell, and Kumar Venkataraman**, “Trading activity and transaction costs in structured credit products,” *Financial Analysts Journal*, 2013, 69(6), 55–68.
- Boudoukh, Jacob, Matthew Richardson, Richard Stanton, and Robert Whitelaw**, “Pricing mortgage-backed securities in a multifactor interest rate environment: A multivariate density estimation approach,” *Review of Financial Studies*, 1997, 10, 405–446.
- Boyarchenko, Nina, Andreas Fuster, and David O. Lucca**, “Understanding Mortgage Spreads,” *The Review of Financial Studies*, 2019, 32 (10), 3799–3850.
- Brennan, Michael J. and Eduardo S. Schwartz**, “Determinants of GNMA Mortgage Prices,” *Real Estate Economics*, 1985, 13 (3), 209–228.
- Chen, Si**, “Understanding Option-Adjusted Spreads,” *The Journal of Portfolio Management*, 1996, 22 (4), 104–113.
- Chernov, Mikhail, Brett R. Dunn, and Francis A. Longstaff**, “Macroeconomic-Driven Prepayment Risk and the Valuation of Mortgage-Backed Securities,” *The Review of Financial Studies*, 12 2017, 31 (3), 1132–1183.
- Cheyette, Oren**, “Implied Prepayments,” *The Journal of Portfolio Management*, 1996, 23 (1), 107–115.

- Chodorow-Reich, Gabriel, Valentin Haddad, and Andra Ghent**, “Asset Insulators,” *Review of Financial Studies*, 2019, *forthcoming*.
- Deng, Yongheng, John M. Quigley, and Robert Van Order**, “Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options,” *Econometrica*, 2000, 68 (2), 275–307.
- Diep, Peter, Andrea L. Eisfeldt, and Scott Richardson**, “The Cross Section of MBS Returns,” *Working Paper*, 2019.
- Downing, Christopher, Dwight Jaffee, and Nancy Wallace**, “Is the Market for Mortgage Backed Securities a Market for Lemons?,” *Review of Financial Studies*, 2009, 22–7, 2457–2494.
- _____, **Richard Stanton, and Nancy Wallace**, “An Empirical Test of a Two-Factor Mortgage Valuation Model: How Much Do House Prices Matter?,” *Real Estate Economics*, 2005, 33 (4), 681–710.
- Duarte, Jefferson**, “The Causal Effect of Mortgage Refinancing on Interest Rate Volatility: Empirical Evidence and Theoretical Implications,” *The Review of Financial Studies*, 12 2007, 21 (4), 1689–1731.
- Duarte, Jefferspm, Francis Longstaff, and Fan Yu**, “Risk and Return in Fixed Income Arbitrage: Nickels in Front of a Steamroller?,” *Review of Financial Studies*, 2007, 20 (3), 769–811.
- Dunn, Kenneth and John McConnell**, “Valuation of GNMA mortgage-backed securities,” *Journal of Finance*, 1981, 36, 599–616.
- Dunn, Kenneth B. and Chester S. Spatt**, “The Effect of Refinancing Costs and Market Imperfections on the Optimal Call Strategy and the Pricing of Debt Contracts,” *Real Estate Economics*, 2005, 33 (4), 595–617.
- Easley, David and Maureen O’Hara**, “Chapter 17 Microstructure and asset pricing,” in “Financial Markets and Asset Pricing,” Vol. 1 of *Handbook of the Economics of Finance*, Elsevier, 2003, pp. 1021 – 1051.
- Fabozzi, Frank and Steven Mann**, “Handbook of Fixed Income Securities,” *McGraw-Hill; 8 edition*, 2011.
- FRBNY**, “Statement Regarding Treasury Securities, Agency Mortgage-Backed Securities, and Repurchase Agreement Operations,” 2020.
- Friewald, Nils, Rainer Jankowitsch, and Marti G. Subrahmanyam**, “Transparency and Liquidity in the Structured Product Market,” *The Review of Asset Pricing Studies*, 2017, 7 (2), 316–348.
- Frost, Josh, Lorie Logan, Antoine Martin, Patrick McCabe, Fabio Natalucci, and Julie Remache**, “Overnight RRP Operations as a Monetary Policy Tool: Some Design Considerations,” *Finance and Economics Discussion Series*, 2015, 2015-010.
- Gabaix, Xavier, Arvind Krishnamurthy, and Olivier Vigneron**, “Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market,” *Journal of Finance*, April 2007, 62 (2), 557–595.

- Gao, Pengjie, Paul Schultz, and Zhaogang Song**, “Liquidity in a Market for Unique Assets: Specified Pool and TBA Trading in the Mortgage Backed Securities Market,” *Journal of Finance*, 2017, 72-3, 1119–1170.
- _____, _____, and _____, “Trading Methods and Trading Costs for Agency Mortgage Backed Securities,” *Journal of Investment Management*, 2018, 16, 29–46.
- Gay, Gerald and Steven Manaster**, “The Quality Delivery Implicit in Futures Contracts,” *Journal of Financial Economics*, 1984, 13, 353–370.
- Gurkaynak, Refet, Brian Sack, and Jonathan Wright**, “The U.S. Treasury Yield Curve: 1961 to the Present,” *Journal of Monetary Economics*, 2007, 54(8), 2291–2304.
- Hansen, Samuel**, “Mortgage Convexity,” *Journal of Financial Economics*, 2014, 113-2, 270–299.
- Hayre, Lakhbir**, “Salomon Smith Barney Guide to Mortgage-Backed and Asset-Backed Securities,” *Wiley*, 2001.
- He, Zhiguo and Zhaogang Song**, “Agency MBS as Safe Assets,” *working paper*, 2019, *Johns Hopkins Carey Business School*.
- _____, **Bryan Kelly, and Asaf Manela**, “Intermediary asset pricing: New evidence from many asset classes,” *Journal of Financial Economics*, 2017, 126 (1), 1–35.
- _____, **Paymon Khorrami, and Zhaogang Song**, “Commonality in Credit Spread Changes: Dealer Inventory and Intermediary Distress,” Working Paper 26494, National Bureau of Economic Research 2019.
- Hegde, S**, “An empirical analysis of implicit delivery options in the Treasury bond futures contract,” *Journal of Banking and Finance*, 1988, 12, 469–492.
- Hemler, Michael**, “The Quality Delivery Option in Treasury Bond Futures Contracts,” *Journal of Finance*, 1990, 45-5, 1565–1586.
- Himmelberg, Charles, Marty Young, Hui Shan, and Chris Henson**, “The Impact of Fed MBS Purchases on TBA Pool Quality,” *Goldman Sachs*, 2013.
- Hu, Xing, Jun Pan, and Jiang Wang**, “Noise as Information for Illiquidity,” *Journal of Finance*, 2013, 68, 2223–2772.
- Ihrig, Jane, Edward Kim, Cindy M. Vojtech, and Gretchen C. Weinbach**, “How Have Banks Been Managing the Composition of High-Quality Liquid Assets?,” *Federal Reserve Bank of St. Louis Review*, 2019, *Third Quarter*, 177–201.
- Kane, A. and A. Marcus**, “The quality option in the Treasury bond futures market: An empirical assessment,” *Journal of Futures Marketse*, 1986, 6, 231–248.

- Kim, You Suk and Yesol Huh**, “The Real Effects of the Secondary Market Trading Structure: Evidence from the Mortgage Market,” *working paper*, 2019.
- Krishnamurthy, Arvind and Annette Vissing-Jorgensen**, “The Ins and Outs of LSAPs,” *Working paper*, 2013.
- Levin, Alexander and Andrew Davidson**, “Prepayment Risk-and Option-Adjusted Valuation of MBS,” *The Journal of Portfolio Management*, 2005, 31 (4), 73–85.
- Li, Wei and Zhaogang Song**, “Asset Heterogeneity, Market Fragmentation, and Quasi-Consolidated Trading,” *working paper*, 2020.
- Longstaff, Francis A.**, “Borrower Credit and the Valuation of Mortgage-Backed Securities,” *Real Estate Economics*, 2005, 33 (4), 619–661.
- Malkhozov, Aytok, Philippe Mueller, Andrea Vedolin, and Gyuri Venter**, “Mortgage Risk and the Yield Curve,” *The Review of Financial Studies*, 01 2016, 29 (5), 1220–1253.
- Richard, Scott F. and Richard Roll**, “Prepayments on fixed-rate mortgage-backed securities,” *The Journal of Portfolio Management*, 1989, 15 (3), 73–82.
- Schultz, Paul and Zhaogang Song**, “Transparency and Dealer Networks: Evidence from the Initiation of Post-Trade Reporting in the Mortgage Backed Security Market,” *Journal of Financial Economics*, 2019, 133, 113–133.
- Schwartz, Eduardo and Walter Torous**, “Prepayment and the valuation of mortgage-backed securities,” *Journal of Finance*, 1989, 44, 375–392.
- Song, Zhaogang and Haoxiang Zhu**, “Quantitative easing auctions of Treasury bonds,” *Journal of Financial Economics*, 2018, 128 (1), 103 – 124.
- _____ and _____, “Mortgage Dollar Roll,” *The Review of Financial Studies*, 2019, 32 (8), 2955–2996.
- Spatt, Chester**, “Frictions in the Bond Market,” *Keynote Speech: Second MTS Conference on Financial Markets*, 2004.
- Stanton, Richard**, “Rational prepayment and the valuation of mortgage-backed securities,” *Review of Financial Studies*, 1995, 8, 677–708.
- _____ and Nancy Wallace, “Mortgage Choice: What’s the Point?,” *Real Estate Economics*, 1998, 26 (2), 173–205.
- Svensson, Lars**, “Estimating and Interpreting Forward Rates: Sweden 1992-1994,” *National Bureau of Economic Research Working Paper #4871*, 1994.
- Titman, Sheridan and Walter Torous**, “Valuing Commercial Mortgages: An Empirical Investigation of the Contingent-Claims Approach to Pricing Risky Debt,” *Journal of Finance*, 1989, 44 (2), 345–373.

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

Vickery, James and Joshua Wright, “TBA Trading and Liquidity in the Agency MBS Market,” *Federal Reserve Bank of New York Economic Policy Review*, 2013, 19.