Understanding Mortgage Spreads
Nina Boyarchenko, Andreas Fuster, and David O. Lucca
Federal Reserve Bank of New York Staff Reports, no. 674
May 2014; revised April 2017
JEL classification: G10, G12, G13

Abstract
Most mortgages in the United States are securitized in agency mortgage-backed securities (MBS), and as a result, yield spreads on these securities are a key determinant of homeowners’ funding costs. We study variation in MBS spreads over time and across securities, and document a cross-sectional “smile” pattern in MBS spreads with respect to the securities’ coupon rates. We propose non-interest-rate prepayment risk as a candidate driver of MBS spread variation and present a new pricing model that uses “stripped” MBS prices to identify the contribution of this prepayment risk to the spread. The pricing model finds that the smile can be explained by prepayment risk, while the time-series variation is mostly accounted for by a non-prepayment risk factor that co-moves with MBS supply and credit risk in other fixed-income markets. We use the pricing model to study the MBS market response to the Federal Reserve’s large-scale asset purchase program and to interpret the post-announcement divergence of spreads across MBS.

Key words: agency mortgage-backed securities, option-adjusted spreads, prepayment risk, OAS smile
“Whoever bought the bonds […] couldn’t be certain how long the loan lasted. If an entire neighborhood moved (paying off its mortgages), the bondholder, who had thought he owned a thirty-year mortgage bond, found himself sitting on a pile of cash instead. More likely, interest rates fell, and the entire neighborhood refinanced its thirty-year fixed rate mortgages at the lower rates. […] In other words, money invested in mortgage bonds is normally returned at the worst possible time for the lender.” — Michael Lewis, Liar’s Poker, Chapter 5

1 Introduction

At the peak of the financial crisis in the fall of 2008, spreads on residential mortgage-backed securities (MBS) guaranteed by U.S. government-sponsored enterprises Fannie Mae and Freddie Mac and the government agency Ginnie Mae spiked to historical highs. In response, the Federal Reserve announced that it would purchase these securities in large quantities to “reduce the cost and increase the availability of credit for the purchase of houses.”1 Mortgage rates for U.S. homeowners reflect movements in MBS spreads as most mortgage loans are securitized in MBS. Following the Fed announcement, spreads on lower-coupon MBS declined sharply, consistent with the program’s objective; at the same time, spreads on higher-coupon MBS widened. This differential spread response suggests that the cross section of MBS prices reflects either compensation for multiple sources of risk, heterogeneous exposures, or both. This paper first characterizes the time-series and cross-sectional variation of MBS spreads on a long sample, and then presents a method to disentangle the contribution of different risk factors to MBS spread variation in this sample and around the Fed announcement.

Credit risk of MBS is limited because of the explicit (for Ginnie Mae) or implicit (for Fannie Mae and Freddie Mac) guarantee by the U.S. government. However, MBS investors are uniquely exposed to uncertainty about the timing of cash flows, as exemplified by the quote above. U.S. mortgage borrowers can prepay the loan balance at any time without penalty, and do so especially as rates drop. The price appreciation from rate declines is thus limited as MBS investors are short borrowers’ prepayment option. Yields on MBS exceed those on Treasuries or interest rate swaps to compensate investors for this optionality. But even after accounting for the option cost associated with interest rate variability, the remaining option-adjusted spread (OAS) can be substantial. Since the OAS is equal to a weighted average of future expected excess returns after hedging for interest

1http://www.federalreserve.gov/newsevents/press/monetary/20081125b.htm. The term “MBS” in this paper refers only to securities issued by Freddie Mac and Fannie Mae or guaranteed by Ginnie Mae (often called “agency MBS”) and backed by residential properties; according to SIFMA, as of 2013:Q4 agency MBS totaled about $6 trillion in principal outstanding. Other securitized assets backed by real estate property include “private-label” residential MBS issued by private firms (and backed by subprime, Alt-A, or jumbo loans), as well as commercial MBS.
rate risk, non-zero OAS suggests that MBS prices reflect compensation for additional sources of risk. We decompose these residuals into risks related to shifts in prepayments that are not driven by interest rates alone, and a component related to non-prepayment risk factors such as liquidity.

To measure risk premia in MBS, we construct an OAS measure based on surveys of investors’ prepayment expectations and also study spreads collected from six different dealers over a period of 15 years. In both cases, we find that, in the time series, the OAS (to swaps) on a market value-weighted index is typically close to zero but reaches high levels in periods of market stress, such as 1998 (around the failure of the Long-Term Capital Management fund) or the fall of 2008. We also document important cross-sectional variation in OAS. At any point in time, MBS with different coupons trade in the market, reflecting disparate rates for mortgages underlying each security. We group MBS according to their “moneyness,” or the difference between the rate on the loans in the MBS and current mortgage rates, which is a key distinguishing feature as it determines borrowers’ incentive to prepay their loans. In this cross section we uncover an “OAS smile”: spreads tend to be lowest for securities for which the prepayment option is at-the-money (ATM), and increase if the option moves out-of-the-money (OTM) or in-the-money (ITM). A similar smile pattern also holds in hedged MBS returns. Correspondingly, a pure long strategy in deeply ITM MBS earns a Sharpe ratio of about 1.9 in our sample, as compared to about 0.7 for a long-ATM strategy. We also show that OAS predict future realized returns, and that realized returns are related to movements in moneyness in a way consistent with the OAS smile.

The OAS smile suggests that investors in MBS earn risk compensation for factors other than interest rates; in particular, these may include other important drivers of prepayments, such as house prices, underwriting standards, and government policies. Variability in these non-interest-rate prepayment factors is not easily diversifiable and will be reflected in the OAS. While the OAS accounts for the predicted path of the non-interest-rate factors, it does not reflect their associated risk premia, because prepayments are projected under the physical, rather than the risk-neutral, measure for these factors. These risk premia, which we refer to as “prepayment risk premia”, cannot be directly measured because market instruments priced off each of these individual factors do not exist.2 Discrete changes in prepayment speeds may occur and are also not easily diversifiable; the associated event risk premium would also be reflected in the OAS.

While prepayment risk premia may give rise to the OAS smile, risk factors unrelated to pre-

---

2Importantly, in our usage, “prepayment risk” does not reflect prepayment variation due to interest rates; instead it is the risk of over- or underpredicting prepayments for given rates.
payment, such as liquidity, could also lead to such a pattern. For example, newly issued MBS, which are ATM and more heavily traded, could command a lower OAS due to better liquidity. Without strong assumptions on the liquidity component, prices of standard MBS (which pass through both principal and interest payments) are insufficient to isolate prepayment risk premia in the OAS. Instead, we propose a new approach based on “stripped” MBS that pass through only interest payments (an “IO” strip) or principal payments (a “PO” strip). The additional information provided by separate prices for these strips on a given loan pool, together with the assumption that a pair of strips is fairly valued relative to each other, allows us to identify market-implied risk-neutral (“Q”) prepayment rates as multiples of physical (“P”) ones. We refer to the remaining OAS when using the Q-prepayment rates as $OAS^Q$, while the difference between the standard OAS and $OAS^Q$ measures a security’s prepayment risk premium.

Our pricing model finds that the OAS smile is explained by higher prepayment risk premia for securities that are OTM and, especially, ITM. There is little evidence that liquidity or other non-prepayment risks vary significantly with moneyness, except perhaps for the most deeply ITM securities. In the time series, instead, we document that much of the OAS variation on a value-weighted index is driven by the $OAS^Q$ component. We show that $OAS^Q$ on the index is related to spreads on other agency debt securities, which may reflect shared risk factors such as changes in the implicit government guarantee or liquidity. Even after controlling for agency debt spreads, OAS are strongly correlated with credit spreads (Baa-Aaa). Given the different sources of risk in the two markets, this finding may suggest the existence of a common marginal investor in corporates and MBS that exhibits time-varying risk aversion, such as an intermediary subject to time-varying risk constraints (for example, Shleifer and Vishny, 1997; Duffie, 2010; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014). Consistently, we find that a measure of supply of MBS (based on new issuance) is also positively related to $OAS^Q$.

The response of OAS to the Fed’s large-scale asset purchases (LSAPs) announcement in November 2008 provides further evidence on the potential role of balance sheet capacity of financial intermediaries in the MBS market. After the announcement, the $OAS^Q$ fell across coupons, as investors anticipated that much of the near-term MBS supply would be held on the Fed’s balance sheet rather than by constrained private investors. However, total OAS diverged, with lower coupon OAS falling and spreads on higher coupons moderately increasing. According to our pricing model, the decline in $OAS^Q$ for higher-coupon MBS is offset by an increase in prepayment risk premia as these securities moved further in the money. In other words, the heterogeneous
OAS response is the direct manifestation of the smile pattern in the prepayment risk premium component in the OAS that this paper emphasizes.

Related literature. Several papers have studied the interaction of interest rate risk between MBS and other markets. This literature finds that investors’ need to hedge MBS convexity risk may explain significant variation in interest rate volatility and excess returns on Treasuries (Duarte, 2008; Hanson, 2014; Malkhozov et al., 2016; Perli and Sack, 2003). Our analysis is complementary to this work as we focus on MBS-specific risks and how they respond to changes in other fixed income markets. More closely related to this paper, Boudoukh et al. (1997) suggest that prepayment-related risks are a likely candidate for the component of MBS prices unexplained by the variation in the interest rate level and slope. Carlin et al. (2014) use long-run prepayment projections from surveys, which we also employ, to study the role of disagreement in MBS returns and their volatility.\(^3\)

Gabaix et al. (2007) study OAS on IO strips from a dealer model between 1993 and 1998, and document that these spreads covary with the moneyness of the market, a fact that they show to be consistent with a prepayment risk premium and the existence of specialized MBS investors. Gabaix et al. do not focus on pass-through MBS and, while their conceptual framework successfully explains the OAS patterns of the IOs in their sample, it predicts a linear, rather than a smile-shaped, relation between a pass-through MBS’s OAS and its moneyness, since they assume that securities have a constant loading on a single-factor aggregate prepayment shock. We show that the OAS smile is in fact a result of prepayment risk but of a more general form, while also allowing for liquidity or other non-prepayment risk factors to affect OAS. Similarly to this paper’s empirical pricing model, Levin and Davidson (2005) extract a market-implied prepayment function from the cross section of pass-through securities.\(^4\) Because they assume, however, that the residual risk premia in the OAS are constant across coupons, the OAS smile in their framework can only be explained by prepayment risk and not liquidity. By using additional information from stripped MBS, this paper relaxes this assumption. Furthermore, we provide a characterization

---

\(^3\)Song and Zhu (2016) and Kitsul and Ochoa (2016) study determinants of financing rates implied by MBS dollar rolls, which are generally affected by liquidity, prepayment and adverse selection risks. Dollar rolls are matched purchases/sales of MBS contracts settling in two subsequent months. While implied financing rates partly reflect MBS liquidity, their calculation relies on prepayment rate expectation under the physical measure and therefore should also incorporate prepayment risk premia as discussed in this paper. Furthermore, as Song and Zhu (2016) emphasize, dollar rolls are strongly affected by adverse selection risk.

\(^4\)Arcidiacono et al. (2013) extend their method to more complex structured securities. Cheyette (1996) and Cohler et al. (1997) are earlier practitioner papers proposing that MBS prices can be used to obtain market-implied prepayments.
of spread patterns over a long sample period, present a conceptual framework to rationalize our findings, and study risk premia covariates.

Two interesting papers subsequent to this work also emphasize the importance of prepayment risk for the cross section of MBS. Chernov et al. (2016) estimate parameters of a simple prepayment function from prices on pass-through MBS. Consistent with our results, they find an important role for a credit/liquidity spread (assumed constant in the cross section) in explaining price variation over time. In terms of prepayment risk, their model implies a dominant role for risks related to turnover independent of refinancing incentives, rather than risks related to refinancing activity of ITM borrowers. Diep et al. (2016) study the cross section of realized MBS excess returns. As in this paper, they find evidence of a smile pattern in their pooled data, with ATM pools earning relatively lower excess returns. However, they argue that different conditional patterns of returns exist depending on whether the MBS market as a whole is ITM or OTM, suggesting that prepayment risk premia change sign with market moneyness. As we show in Section 2, the smile pattern in expected excess returns as measured by OAS holds irrespective of market type, and cross-sectional patterns in returns are also consistent with the smile when we study their relation with changes in mortgage rates (and thus moneyness).

From a methodological perspective, this paper is related to studies that confront asset pricing models with both physical and risk-neutral data. Driessen (2005), uses U.S. bond price data and historical default rates to estimate a default event risk premium. He parameterizes the risk-neutral intensity of default as a multiple of the historical intensity; in this paper, we follow a similar approach in parametrizing the risk-neutral prepayment path as a multiple of the prepayment path under the historical measure. Almeida and Philippon (2007) use the default risk premia estimated for bonds with different credit ratings to compute the risk-adjusted costs of financial distress for firms in the same ratings class. Our exercise is similar in spirit as we use the prepayment risk premia estimated for different MBS pools to compute prepayment risk-adjusted liquidity costs faced by investors in this market. Outside of default risk premia, similar approaches have been used in the estimation of interest rate risk premia (Stanton, 1997), variance risk premia (see e.g. Carr and Wu, 2009; Trolle and Schwartz, 2009), and in other asset classes.
2 Facts about mortgage spreads

In this section we provide a brief overview of MBS and define the option-adjusted spread (OAS). We then characterize time-series and cross-sectional spread variation in terms of a few stylized facts, and relate the OAS to MBS returns.

2.1 The agency MBS market and spreads

In an agency securitization, a mortgage originator pools loans and then delivers the pool in exchange for an MBS certificate, which can be subsequently sold to investors in the secondary market.5 Servicers, which are often affiliated with the loan originator, collect payments from homeowners that are passed on to MBS holders after deducting a servicing fee and the agency guarantee fee. In a standard MBS, also known as a pass-through, homeowners’ payments (interest and principal) are assigned pro-rata to all investors. However, cash flow assignment rules can be more complicated with multiple tranches, as is the case for stripped MBS. We focus on MBS backed by fixed-rate mortgages (FRMs) with original maturities of 30 years on 1-4 family properties; these securities account for more than two-thirds of all agency MBS.6

In agency MBS, the risk of default of the underlying mortgages is not borne by investors but by the agencies that guarantee timely repayment of principal and interest. Because of this guarantee, agency MBS are generally perceived as being free of credit risk. But while Ginnie Mae securities have the full faith and credit of the U.S. government, assessing credit risk of Fannie Mae and Freddie Mac securities is more complex. Government backing for these securities is only implicit and results from investors’ anticipation of government support under a severe stress scenario, as when Fannie Mae and Freddie Mac were placed in federal conservatorships in September 2008.7

Beyond the implicit guarantee, a distinct feature of MBS is the embedded prepayment option: borrowers can prepay their loan balance at par at any time, without paying a fee. Because borrowers are more likely to do so when rates decline, MBS investors are exposed to reinvestment risk and have limited upside as rates decline; more formally, they are short an American option. The embedded prepayment option is crucial in the valuation of MBS, since it creates uncertainty.

---

5 In addition to these “lender swap” transactions, Fannie Mae and Freddie Mac also conduct “whole loan conduit” transactions, in which the agencies buy loans against cash from (typically smaller) originators, pool these loans themselves, and then market the issued MBS.
6 As of March 2014, the balance-weighted share was 69 percent (author calculations based on data from eMBS).
7 The conservatorships have resulted in an effective government guarantee of Fannie and Freddie securities since September 2008, but (at least in principle) this guarantee is still temporary and thus not as strong as the one underlying Ginnie Mae securities.
in the timing of future cash flows, $X_t$. Prepayment rates depend on loan characteristics as well as macroeconomic factors such as house prices and variation in interest rates, which are the key determinant of refinancing activity.\footnote{Appendix A provides a detailed description of how MBS cash flows depend on prepayments and scheduled amortization.} The model price of an MBS is equal to the average discounted value of possible cash-flow paths, and the OAS is defined as the constant spread to the path of discount rates $r_{t+j}$ that equates model and market prices (see Hayre, 2001, for example):

$$P_t = E_{Q_r}^Q \left[ \sum_{k=1}^{T-t} \frac{X_{t+k}}{\prod_{j=1}^{k} (1 + r_{t+j} + \text{OAS})} \right], \quad (2.1)$$

where $P_t$ the market price of an MBS, and $E_{Q_r}^Q$ is the expectation under the interest-rate-risk-neutral measure $Q_r$. The OAS increases the greater the discounted value of cash flows relative to the market price, meaning that an MBS trading below the model price after accounting for the prepayment option will have a positive OAS.\footnote{In Appendix H, we study spreads that do not account for interest rate volatility (often called zero-volatility spreads, or ZVS) as well as the difference between these spreads and the OAS, which is a measure of the value of the embedded prepayment option.}

If interest rates were the only factor affecting prepayments, one would expect the OAS to be equal to zero. However, given that non-interest-rate factors such as house prices or lending standards also impact household prepayment activity, the OAS will reflect compensation for systematic risk associated with these other factors. This is because the OAS is calculated under the risk-neutral measure for interest rates but the physical measure for prepayments (as a function of rates). The reason why the OAS does not incorporate non-interest-rate factors is that market instruments priced off each of these factors do not exist; thus, risk premia associated with these factors cannot be measured directly. As a result, risk premia attached to these factors’ innovations are reflected in the OAS. In addition, the OAS may also reflect MBS liquidity discounts and a prepayment event risk premium resulting from unhedgeable risk due to discrete changes in prepayment speeds. Thus, one should not expect the OAS to equal zero, and the goal of this paper is to understand the sources of its variation.

We next characterize OAS variation in the MBS universe using a market value-weighted index (in the time series) as well as in terms of MBS moneyness (in the cross section). Whenever possible we consider in our analysis OAS relative to swaps, rather than Treasuries, since these instruments are more commonly used for hedging MBS (see e.g. the discussion in Duarte, 2008) and also because interest rate volatility measures, used to calibrate the term structure model, are more...
readily available for swaps. Throughout the paper, OAS are expressed in basis points per year, following market convention. We use spreads in the to-be-announced (TBA) market, where the bulk of MBS trading happens. The TBA market is a forward market for pass-through MBS where a seller and buyer agree on a select number of characteristics of the securities to be delivered (issuer, maturity, coupon, par amount), a transaction price, and a settlement date either 1, 2, or 3 months in the future. The precise securities that are delivered are only announced 48 hours prior to settlement, and delivery occurs on a “cheapest-to-deliver” basis (see Vickery and Wright, 2013, for a detailed discussion). Because OAS are model-dependent, we collected end-of-month OAS on Fannie Mae securities from six different dealers over the period 1996 to 2010. As a result, the stylized facts we present are robust to idiosyncratic modeling choices of any particular dealer and, through data-quality filters we impose, issues arising from incorrect or stale price quotes. Further details on the sample, the data-quality filters, and descriptive statistics are provided in Appendix B.

### 2.2 OAS variation in the time series

At each point in time, MBS with different coupons coexist, primarily as a result of disparate loan rates of the mortgages underlying an MBS. The benchmark contract in the TBA market is the so-called current coupon, a synthetic 30-year fixed-rate MBS obtained by interpolating the highest coupon below par and the lowest coupon above par. The interest in this benchmark is due to the fact that most newly originated mortgages are securitized in coupons trading close to par. Despite its benchmark status, the current coupon is not representative of the MBS universe as a whole, because at any point in time, only a limited fraction of the universe is in coupons trading close to par. This is illustrated in Figure 1. For example, the current coupon at the end of 2010 was around 4 percent (red line, measured on the right y-axis) but securities with a coupon of 4

---

10 Feldhütter and Lando (2008) study the determinants of spreads between swaps and Treasuries and find that they are mostly driven by the convenience yield of Treasuries, though MBS hedging activity may also play a role at times.

11 Freddie Mac securities are generally priced relatively close to Fannie Mae’s, reflecting the similar collateral and implicit government backing. The prices of Ginnie Mae securities can differ significantly (for the same coupon) from Fannie and Freddie MBS, reflecting the difference in prepayment characteristics (Ginnie Mae MBS are backed by FHA/VA loans) and perhaps the explicit government guarantee. Throughout this paper, we focus on Fannie Mae MBS.

12 These differences arise due to variation in loan origination dates, as well as other factors such as “points” paid (or received) by the borrowers. By paying points at origination (with one point corresponding to one percent of the loan amount), borrowers can lower the interest rate on their loan. Conversely, by accepting a higher rate, borrowers receive a “rebate” that they can use to cover origination costs.

13 Alternatively, it is obtained by extrapolating from the lowest coupon above par in case no coupon is trading below par (which has frequently been the case in recent years). Sometimes the term “current coupon” is used for the actual coupon trading just above par; we prefer the term “production coupon” to refer to that security.
percent accounted for only about 20 percent of the total outstanding on a balance weighted basis. Another limitation of the current coupon is that since it is a synthetic contract, variation in its yield or spreads can be noisy because of inter- and extrapolations from other contracts and the required assumptions about the characteristics of loans that would be delivered in a pool trading at par (see Fuster et al., 2013, for more detail).

To characterize time-series OAS variation, we therefore follow the methodology of fixed income indices (such as Barclays and Citi, which are main benchmarks for money managers) and construct a market value-weighted index (the “TBA index”) based on the universe of outstanding pass-through MBS. In contrast to other indices we do not rely on any particular dealer’s pricing model; instead, we average the OAS for a coupon across the dealers for which we have quotes on a given day, and then compute averages across coupons using the market value of the remaining principal balance of each coupon in the MBS universe.

The resulting time series of spreads on the TBA index is shown in the top panel of Figure 2. The OAS on the value-weighted index is typically close to zero, consistent with the view that credit risk of MBS is generally limited; however, the OAS spiked to more than 150 basis points in the fall of 2008, and also rose significantly around the 1998 demise of the Long-Term Capital Management fund. To provide initial evidence on potential drivers of this time-series variation we study the relation between the OAS on the TBA index and commonly used fixed income risk factors. Table 1 shows estimated coefficients from a regression of the OAS on: (i) the convenience yield on Treasury securities (reflecting their liquidity and safety) as measured by the Aaa-Treasury spread; (ii) credit spreads as measured by the Baa-Aaa spread; (iii) the slope of the yield curve (measured by the yield difference between 10-year Treasury bonds and 3-month Treasury bills); and (iv) the swaption-implied volatility of interest rates.\(^{14}\) The OAS on the TBA index is strongly related to credit spreads (and to a lesser extent to the Aaa-Treasury spread) both over the full sample (column 1) and the pre-crisis period (ending in July 2007, column 2), and is largely unaffected by the other risk measures. This suggests the existence of common pricing factors between the MBS and corporate bond markets.\(^{15}\) In contrast, implied rate volatility does not explain the OAS variation,

\(^{14}\) All right-hand-side variables are standardized so that each coefficient estimate can be interpreted as the spread impact in basis points of a unit standard deviation increase. As in Krishnamurthy and Vissing-Jorgensen (2012) the Aaa-Treasury spread is the difference between the Moody’s Seasoned Aaa corporate bond yield and the 20-year constant maturity Treasury (CMT) rate. The Baa rate is also from Moody’s, and bill rates and 10-year Treasury yields are CMTs as well. All rates were obtained from the H.15 release. Swaption quotes are basis point, or normal, volatility of 2-year into 10-year contracts, from JP Morgan. We choose the Newey-West lag length based on the Stock-Watson rule-of-thumb measure \(0.75 \times T^{1/3}\).

\(^{15}\) Brown (1999) relates the OAS to Treasuries of pass-through MBS over the period 1993–1997 to other risk premia
which is to be expected since the OAS adjusts for interest rate risk and thus should not reflect interest rate uncertainty. The slope of the yield curve, often used as a proxy for term premia, is also not systematically related to the OAS. In Section 5, we return to the determinants of the time-series variation in spreads, focusing on mortgage-specific risk factors.

2.3 OAS variation in the cross section: the OAS smile

While variation in OAS on the TBA index is informative of the MBS market as a whole, it masks significant variation in the cross section of securities. As discussed above, this cross section is composed of MBS with different underlying loan rates. This rate variation across MBS leads to borrower heterogeneity in their monetary incentives to refinance. We refer to this incentive as a security’s “moneyness” and define it (for security \( j \) at time \( t \)) as

\[
\text{Moneyness}_{j,t} = \text{Coupon}_j + 0.5 - \text{FRMrate}_t,
\]

where \( \text{FRMrate}_t \) is the mortgage rate on new loans at time \( t \), measured by the end-of-month value of the Freddie Mac Primary Mortgage Market Survey rate on 30-year FRMs. We add 0.5 to the coupon rate because the mortgage loan rates are typically around 50 basis points higher than the MBS coupon.\(^{16}\) When moneyness is positive, a borrower can lower his monthly payment by refinancing the loan—the borrower’s prepayment option is “in-the-money” (ITM)—while if moneyness is negative, refinancing (or selling the home and buying another home with a new mortgage of equal size) would increase the monthly payment—the borrower’s option is “out-of-the-money” (OTM). Aside from determining the refinancing propensity of a loan, moneyness also measures an investor’s gains or losses (in terms of coupon payments) if a mortgage underlying the security prepayrs (at par) and he reinvests the proceeds in a “typical” newly originated MBS (which will approximately have a coupon equal to \( \text{FRMrate}_t \) minus 50 basis points).

The bottom panel of Figure 2 shows the (pooled) variation of spreads as a function of security moneyness. OAS display a smile-shaped pattern: they are lowest for at-the-money (ATM) securities and increase moving away in either direction, especially ITM. OAS on deeply ITM securities and finds a significant correlation of the OAS with spreads of corporate bonds to Treasuries. He interprets his findings as implying a correlation between the market prices of credit risk or liquidity risk on corporates and that of prepayment risk on MBS, but notes that it could also be driven by time variation in the liquidity premium on Treasuries. \(^{16}\)The difference gets allocated to the agency guarantee fee as well as servicing fees (see Fuster et al., 2013, for details). We could alternatively use a security’s “weighted average coupon” (WAC) directly, but the WAC is not known exactly for the TBA securities studied in this section.
on average exceed those on ATM securities by 50 basis points or more.

In Table 2 we report results from regressions that allow us to more precisely quantify this pattern and assess its statistical significance. Instead of imposing parametric restrictions, we simply regress OAS on 50-basis-point moneyness bin dummies, with \([-0.25, 0.25]\) as the omitted category. Column 1 of the top panel studies the pooled variation over the full sample (as displayed in Figure 2), confirming a statistically and economically significant smile pattern. Column 2 shows that this pattern remains almost unchanged if we control for month fixed effects (meaning that only cross-sectional variation is exploited). In columns 3 and 4, we repeat the same regressions but end the sample before the onset of the financial crisis in August 2007. For ITM coupons, this somewhat reduces the relative spread differences to ATM securities, but the differences remain monotonic in absolute moneyness and highly statistically significant.

In panel (b) of the table, we split the sample in terms of the moneyness of the agency MBS market as a whole, in order to investigate changes in cross-sectional OAS patterns. As we discuss in more detail in Section 3, the model in Gabaix et al. (2007) predicts that the relationship of OAS with moneyness is linear at a point in time, but can be either upward sloping (if the market is ITM) or downward sloping (if the market is OTM). More recently Diep et al. (2016) present a similar theory which also predicts that prepayment risk premia change sign with market moneyness. Instead, our estimates indicate that the OAS smile is present irrespective of whether the market is ITM or not, even though coefficient estimates vary somewhat and are not always very precise due to limited number of observations (e.g. for ITM bins in OTM markets). In Appendix C, we present further robustness evidence on the smile pattern in OAS, showing for instance that it holds when excluding outliers or coupons with low remaining balance.

2.4 OAS and MBS returns

The OAS is a valuation measure that is widely tracked by financial market participants but that has also been called into question for its model dependence (Kupiec and Kah, 1999). We discuss this issue in the context of the relation between the OAS and MBS returns. We first derive expressions for OAS in terms of returns, test these in the data and finally show implications of the OAS smile for the relation between MBS returns and changes in interest rate.

Shiller, Campbell, and Schoenholtz (1983) derive first-order approximations of the yield to maturity on coupon-bearing Treasury security in terms of future bond returns and of contemporaneous returns and yield changes. We provide two analogous expressions linking the OAS to MBS
returns. Despite the analogy, our derivations (shown in Appendix D) differ from theirs because MBS cash flows $X_t$ are uncertain. Let $rx_{t+1}$ be the one-period MBS excess return:

$$1 + r_{t+1} + rx_{t+1} = \frac{P_{t+1} + X_{t+1}}{P_t},$$

where $r_{t+1}$ is the one-period risk-free rate. Taking expectations of both sides under the interest-rate-risk-neutral measure $Q_r$, and iterating forward we obtain:

$$P_t = E^{Q_r}_t \left[ \sum_{k=1}^{T-t} \frac{X_{t+k}}{\prod_{j=1}^{k} \left( 1 + r_{t+j} + E^{Q_r}_{t+j-1} [rx_{t+j}] \right)} \right]. \quad (2.2)$$

Equating $P_t$ from the last expression to (2.1), linearizing around zero OAS and excess returns, and solving for OAS, yields:

$$OAS_t \approx E^{Q_r}_t \left[ \sum_{k=1}^{T-t} w_{kt} E^{Q_r}_{t+k-1} [rx_{t+k}] \right], \quad (2.3)$$

where the weights $w_{kt}$ (reported in Appendix D) decline in horizon $k$. In words, the OAS is a weighted average of expected one-period excess returns over the lifetime of the security under the interest-rate-risk-neutral measure $Q_r$. A second relation between OAS and returns can be derived by linearizing (2.2) for small changes in OAS$_t$:

$$rx_{t+1} \approx OAS_t - D_t (OAS_{t+1} - OAS_t). \quad (2.4)$$

The realized excess return on an MBS is approximately equal to the sum of spread income (or carry) and the capital gain/loss resulting from changes in spreads, which are greater the larger the MBS duration, or sensitivity, with respect to its spread ($D_t$).

Breeden (1994) shows that the OAS predicts future MBS excess returns between 1988 and 1994, consistent with (2.3). In the remainder of this section, we extend this result to a longer sample, then test the contemporaneous relation between excess returns and OAS implied by (2.4), and finally show that excess returns also exhibit a smile pattern with respect to moneyness, a pattern that is also evident in how the cross section of MBS returns depends on changes in interest rates.

Constructing MBS returns is complex because of the large number of securities, different pricing conventions, and security-specific prepayments. We rely on monthly return data from the MBS sub-components by coupon of the Bloomberg Barclays Aggregate Bond Index, which is the
leading benchmark for fixed income index funds. MBS returns are available both unhedged (that is, as measured from MBS prices and prepayments alone) and interest-rate-hedged (relative to a duration-matched portfolio of Treasury securities).\footnote{The return on an MBS is equal to the sum of price appreciation, coupon yield and paydown return. Because more seasoned securities often trade at a premium to the TBA price, the Barclays index adjusts the capital gain return component with “payup” information. In addition, the calculation of the index incorporates cusip-specific prepayment information to compute the paydown return. The index uses same-day settlement as opposed to standard PSA settlement (fixed monthly dates) as it is typically the case in TBA trading, which is associated with discrete price drops on PSA dates, which is when the attribution of prepayments and coupons is determined (see Chapter 29 in Fabozzi, 2016, for more detail). The excess return for an MBS is calculated as the difference between its total return and that of the equivalent Treasury position, where the equivalent position is obtained from key-rate durations, which are sensitivities to the movement of specific parts of the yield curve. For additional information see, for example, Lehman Brothers (2008). We have also verified that qualitative patterns are similar when using data from a different dealer (results available upon request).} We approximate the $Q_r$ interest rate measure expectation of excess returns with expected hedged returns, thus taking out the component of expected excess returns earned due to interest rate uncertainty, but also analyze unhedged returns for robustness. Bloomberg Barclays MBS returns have recently been analyzed by Diep, Eisfeldt, and Richardson (2016), and we match their 1994-mid 2016 sample period and size cutoff (excluding coupons with less than $1$ billion in outstanding principal). Unlike in the previous subsections, here we rely on the Barclays OAS relative to Treasuries that covers this entire sample period.

From (2.3), the OAS is approximately equal to the expectation under $Q_r$ of the weighted future excess returns. As we show in Appendix D, expected excess hedged returns under $Q_r$ coincide with expected excess hedged returns under the physical measure $P$, which we measure in the data. The first two columns of panel (a) of Table 3 report estimates of a regression of future 1-year hedged MBS returns $(t \rightarrow t + 12)$ on OAS$_t$ either including or excluding time fixed effects; panel (b) shows corresponding coefficients for unhedged returns. Estimated loadings on the OAS range between 0.8 and 1.5 and are highly statistically significant (based on Newey-West standard errors with 18 lags).\footnote{As it is well known in the literature (Hansen-Hodrick), the overlapping return sample generates an MA(12) component in the error term. We use NW with 18 lags to guarantee a positive definite variance covariance matrix and counteract the underweighting of higher covariances from the NW kernel function. Realized returns in the sample move about one-for-one with the OAS; we do not have a sharp prediction on the size of the coefficient since the right-hand side of (2.3) features declining weights whereas in our empirical implementation we apply equal weights and truncate at one year.} Columns 3 and 4 repeat the same exercise using 1-month returns. The loading of excess returns on the lagged OAS is about 1.8 both with and without time fixed effects for hedged returns ($p < 0.01$ based on standard errors clustered by month), and it is similar for unhedged returns. In sum, OAS predict realized hedged and unhedged returns at the 1-month and 1-year horizon consistent with the prediction of (2.3).

To test (2.4), we extend the 1-month regression to include contemporaneous changes in the
OAS. As predicted, the coefficient on the spread change is always negative \((p < 0.01)\) across the four specifications with point estimates that range between -1.8 and -3.3. The adjusted \(R^2\) in the hedged return regression when omitting time effects exceeds 50%, meaning that changes in OAS explain much of the realized variation in hedged returns.

A key feature of the cross section of OAS is the smile pattern with respect to a security’s moneyness (see Section 2.3). Figure 3 shows that 1-month excess returns display a similar pattern, which is also reflected in differential Sharpe ratios (SRs) in the MBS cross section. We compute (annualized) SRs based on monthly Barclays index returns for portfolios of ITM, ATM and OTM MBS securities. Relative to ATM securities, we find much larger SRs for non-ATM securities: using unhedged returns (minus the risk-free rate from Ken French’s website), the SR of a long-ITM (long-OTM) portfolio is 1.86 (1.15) compared to a SR of 0.66 for a long-ATM position. Looking at hedged returns, we find a similar pattern, although it is more pronounced for ITM securities (ITM SR = 0.68; OTM SR = 0.12; ATM SR = 0.04).

The above analysis established that there is a tight link between realized hedged returns and changes in OAS: when OAS fall, realized returns tend to be high. In addition, we have documented a smile-shaped pattern both in expected and realized returns. We now combine these relations to test an additional prediction about the link between realized returns and changes in mortgage rates, without relying on OAS directly. Movements in mortgage rates change the moneyness of MBS by moving them along the smile, and the OAS smile predicts an opposite effect of rate changes on hedged returns depending on whether an MBS is ITM or not. For an OTM MBS, changes in OAS and changes in rates are positively related, as the MBS moves closer to being ATM when rates fall. This implies a negative relationship between rate changes and contemporaneous hedged returns on OTM securities. Conversely, for ITM securities, hedged returns and rate changes should be positively related.

Panel (c) of Table 3 shows that as predicted, hedged returns are negatively (positively) related to changes in mortgage rates for OTM (ITM) securities, irrespective of whether fixed effects are included or not (columns 1 and 2). In columns 3 and 4, we re-estimate this relation in a split sample based on whether the market as a whole is ITM or not. This allows us to test whether

---

19 We compute the ATM portfolio using the return on the coupon that is closest to zero moneyness in each period, as long as it is not more than 25 basis points ITM or OTM. For the ITM (OTM) portfolio we use the most ITM (OTM) coupon as long as it is at least 100 basis points ITM (OTM).

20 The OAS smile implies that OTM securities tend to outperform ITM securities when rates fall, as for example was the case following the November 2008 LSAP announcement, discussed in Section 5.3.

21 When we add month fixed effects, we can only test whether the relationship between returns and rate changes is more positive for ITM securities; the uninteracted coefficient on the mortgage rate change is not identified.
prepayment risk premia change sign as a function of the market overall moneyness as predicted by Gabaix et al. (2007) and Diep et al. (2016). Contrary to what is implied by these theories, but consistent with our OAS smile evidence in Section 2.3, we do not find that the relationship between returns and rate changes flips sign with market moneyness. We return to this issue in the next section when we present the conceptual framework to analyze the OAS smile.

In sum, our analysis of MBS returns shows that, while model-dependent, the OAS is related to realized returns, and that a smile pattern is also evident in the cross section of hedged returns. In the remainder of the paper, we focus on the OAS rather than realized returns, since it is a more direct and less noisy measure of expected excess returns. Also, as the last part of our analysis above illustrated, realized returns across securities are affected by changes in realized rates, which in finite sample could lead to average realized and expected returns having different patterns.

3 Conceptual framework

In this section, we discuss a simple conceptual framework that can be used to understand the sources of risk premia in the valuation of MBS. Here, we illustrate the main intuition behind our decomposition of OAS into a prepayment risk premium and a liquidity risk premium under simple assumptions on the evolution of the interest rate on the loans and the prepayment intensity function, and discuss the derivation under more general assumptions in Appendix E.

Consider a mortgage pool \( j \) with coupon rate \( c_j \) and remaining principal balance \( \theta_{jt} \) at time \( t \). The pool prepays with intensity \( s_{jt} \), so that the principal balance evolves (in continuous time) as

\[
d\theta_{jt} = -s_{jt}\theta_{jt}dt.
\]

We assume that the prepayment rate \( s_{jt} \) is a function of the interest rate incentive \( c_j - r \), where \( r \) is the interest rate on new loans, and a vector of parameters \( \gamma_t = (\gamma_{1t}, \gamma_{2t}, \ldots, \gamma_{Nt})' \), which are uncertain and give rise to non-interest-rate prepayment risk. For ease of notation, we assume here that the interest rate \( r \) is constant, though the derivation in Appendix E shows that the main expression of interest is the same if we assume that rates follow a diffusion. Suppose the prepayment

\[\text{As mentioned earlier, in Gabaix et al. (2007) the relationship of OAS with moneyness is linear and upward sloping if the market is ITM or downward sloping if the market is OTM. This theory thus predicts that in an ITM market, the relationship between hedged returns and mortgage rate changes should be positive for all securities, while in an OTM market, the relationship should be negative for all securities.}\]
rate is given by
\[ s_{jt} = f \left( \gamma_t, c_j - r \right). \]

The parameters \( \gamma_t \) are time-varying, with normal innovations, so that
\[ d\gamma_t = \mu_\gamma dt + \sigma_\gamma dZ_{\gamma t}, \]
where \( Z_{\gamma t} \) is a standard Brownian motion vector. The other source of uncertainty in the model is the liquidity of the securities. We assume that, with intensity \( \mu_t \), the whole market experiences a liquidity event in which a pool \( j \) loses a fraction \( \alpha_{jt} \) of its market value. Thus \( \alpha_{jt} \) should be thought of as how well the security performs in a “bad” market, similar to Acharya and Pedersen (2005). Alternatively, \( \alpha_{jt} \) could be interpreted as the price impact of a decline in the strength of the agency guarantee.

Under no-arbitrage there exists a pricing kernel \( M_t \) such that the time \( t \) price of a future stream of cash flows \( X_{t+s} \) is
\[ P_t = \mathbb{E}_t \left[ \int_0^{\infty} \frac{M_{t+s}}{M_t} X_{t+s} ds \right]. \]
This is analogous to the discrete-time expression for the price of the security (2.1): instead of taking expectations under the interest-rate-neutral measure \( Q_r \) and discounting future cash-flows \( X_t \) using the risk-free rate curve shifted by the OAS, we take expectations under the physical measure and discount using the pricing kernel. Let \( R_t \) be the return to holding a claim to the stream of cash flows \( X_t \), which evolves as
\[ dR_t = \frac{dP_t}{P_t} + \frac{X_t}{P_t} dt. \]
The no-arbitrage restriction (3.1) implies that the expected return can be represented as
\[ \mathbb{E}_t [dR_t] = r dt - \mathbb{E}_t \left[ \frac{dM_t}{M_t} \frac{dP_t}{P_t} \right], \]
where \( r \) is the risk-free rate. As shown in Section 2.4, the OAS is, to a first-order approximation, the lifetime discounted sum of instantaneous risk premia paid to an investor for holding the claim to \( X \). For tractability, in this conceptual framework, we approximate the OAS by the instantaneous
risk premium, denoted \( rp_t \), so that
\[
OAS_t \approx rp_t \equiv -\mathbb{E}_t \left[ \frac{dM_t}{M_t} \frac{dP_t}{P_t} \right].
\]

That is, the risk premium is the covariance between innovations to the price of a security and innovations to the pricing kernel. To solve for the OAS, denote by \( \pi_{\gamma_t} \) the vector of prices of risks associated with innovations to the prepayment model parameters, and \( \pi_{lt} \) the price of risk associated with the liquidity shock. As is standard, these risk prices are given by the co-variation between the innovations to the pricing kernel and the prepayment and liquidity shocks. Since the liquidity risk is undiversifiable for an investor holding a portfolio of MBS, we have \( \pi_{lt} > 1 \) (see e.g. Driessen, 2005); we discuss the sign of \( \pi_{\gamma_t} \) below.

We show in Appendix E that, with these assumptions, the OAS is given by
\[
OAS_{jt} = \alpha_{jt} \mu_t (\pi_{lt} - 1) - \pi'_{\gamma_t} \sigma_{\gamma_t} \frac{1}{P_{jt}} \frac{\partial P_{jt}}{\partial \gamma_t}. \tag{3.2}
\]

Thus, differences in OAS across securities could be the result of (i) differential exposure \( \alpha_j \) to the liquidity shock, or (ii) differential price sensitivity to the prepayment parameters, that is, differential exposure to prepayment risk. Though the expression in eq. (3.2) relies on the particular assumptions for the dynamics of interest rates, prepayment shocks, and liquidity shocks made in this section, we show in Appendix E that this decomposition holds more generally.

We can gain further intuition on potential risk premium variation across securities by taking the first-order approximation of \( P_{jt} \) around the price in the no-uncertainty case,
\[
P_{jt} = 1 + \frac{c_j - r}{r + s_{jt}}, \tag{3.3}
\]
which implies that security \( j \) trades at a premium \( (P_{jt} > 1) \) if \( c_j - r > 0 \) and at a discount \( (P_{jt} < 1) \) if \( c_j - r < 0 \). It also shows that for premium securities the price declines as prepayments \( s_{jt} \) rise, while discount securities benefit from faster prepayments.

Using this and applying the chain rule \( \frac{\partial P}{\partial \gamma} = \frac{\partial P}{\partial s} \cdot \frac{\partial s}{\partial \gamma} \), we get the following approximate expression for the OAS:
\[
OAS_{jt} \approx \alpha_{jt} \mu_t (\pi_{lt} - 1) + \pi'_{\gamma_t} \sigma_{\gamma_t} \frac{c_j - r}{(s_{jt} + c_j) (s_{jt} + r)} \frac{\partial s_{jt}}{\partial \gamma_t}. \tag{3.4}
\]
We can use this expression to understand the conditions under which prepayment risk can lead to the OAS smile. We consider three stylized representations of borrowers’ prepayment behavior. In each case, \( s_j \) corresponds to the expected prepayment speed on security \( j \).

**Case 1:** \( s_{jt} = s_j + \gamma_1 t \beta_j \). This is essentially the framework studied by Gabaix et al. (2007). Each pool has a constant exposure \( \beta_j \) to a single market-wide prepayment shock \( \gamma_1 \). Under this functional form, the \( OAS_{jt} \) in (3.2) will be a linear function of moneyness \( c_j - r \) (regardless of the sign of the risk price \( \pi_{\gamma_1} \)). This case is therefore inconsistent with the OAS smile.

**Case 2:** \( s_{jt} = s_j + \gamma_1 (c_j - r) \). Like in the previous case, a single factor drives prepayment behavior, but the security’s exposure to the prepayment shock now depends on its moneyness, which varies over time.\(^{23}\) This functional form implies that when ITM securities prepay faster than expected (a positive shock to \( \gamma_1 \)), OTM securities prepay slower than expected. This may arise because of mortgage originators’ capacity constraints in (larger than expected) refinancing waves.\(^{24}\) It is easy to see from (3.2) that this case would lead the OAS to be quadratic in \( c_j - r \) (the risk price \( \pi_{\gamma_1} \) is positive in this case, since every security has a positive exposure to \( \gamma_1 \)), and therefore could rationalize the OAS smile.

**Case 3:** \( s_{jt} = s_j + \gamma_1 t \mathbb{1}_{c_j < r} + \gamma_2 t \mathbb{1}_{c_j \geq r} \). In this multi-factor formulation, OTM and ITM prepayments are driven by different shocks (which for simplicity we assume to be orthogonal). For instance, \( \gamma_1 \) might represent the pace of housing turnover while \( \gamma_2 \) might be the effective cost of mortgage refinancing (which varies with underwriting standards and market competitiveness). In equilibrium, the signs of the prices of risk are determined by the average exposure of the representative investor. Holding a portfolio of ITM and OTM securities, this investor will have a negative exposure to \( \gamma_1 \) risk (since OTM securities benefit from fast prepayment) and a positive exposure to \( \gamma_2 \) risk (since the price of ITM securities declines with faster prepayments). Thus \( \pi_{\gamma_1, t} < 0 \) and \( \pi_{\gamma_2, t} > 0 \), resulting in a positive OAS for both ITM and OTM securities and a (v-shaped) OAS smile.

\(^{23}\) \( \gamma_1 > 0 \) because in practice prepayments are a non-decreasing function of the incentive to refinance.  

\(^{24}\) When capacity is tight, mortgage originators may be less willing to originate purchase loans (which are more labor intensive), and they may reduce marketing effort targeted at OTM borrowers (for instance, to induce them to cash out home equity by refinancing their loan). Fuster et al. (2017) show that originators’ pricing margins are strongly correlated with mortgage application volume, consistent with the presence of capacity constraints.
In sum, the OAS displays a smile pattern in moneyness if prepayments are driven by the specification in case 2 or 3, but not in the single-factor representation of case 1. More generally, prepayment risk premia can explain the OAS smile whenever OTM securities are not a hedge for ITM pools (as they would be in case 1).²⁵ To separate liquidity and prepayment risk premia, in the next section we provide a method to identify a “prepayment-risk-neutral OAS,” denoted \( OAS^Q \), as the spread that only reflects liquidity risk:

\[
OAS^Q_{jt} = \alpha_{jt} \mu_t (\pi_{lt} - 1) .
\]  

(3.5)

The prepayment risk premium paid to the investor is then just the difference between the OAS and \( OAS^Q \).

4 Pricing model: Decomposing the OAS

In this section, we propose a method to decompose the “standard” OAS into a prepayment risk premium component and a remaining risk premium (\( OAS^Q \)). We then implement this method using a pricing model, which consists of an interest rate and a prepayment component. In contrast to standard approaches, such as Stanton (1995) or practitioner models, we employ information from stripped MBS to identify a market-implied prepayment function and the contribution of prepayment risk to the OAS.

4.1 Identification of \( OAS^Q \)

As discussed in the Section 2, the OAS only accounts for interest rate uncertainty (and only interest rates are simulated in empirical pricing models) and ignores other sources of prepayment risk, such as uncertainty about house prices or lending standards. In this section we propose a method to identify a risk-neutral prepayment function obtained from market prices, then compute an OAS using this function (\( OAS^Q \)) and finally obtain the contribution of prepayment risk to the OAS.

Following the credit risk literature (e.g. Driessen, 2005), we assume that the market-implied risk-neutral (“Q’) prepayment function is a multiple \( \Lambda \) of the physical (“P’) one. We allow the multiplier \( \Lambda \) to be pool-specific to account for differences in pools’ sensitivities to non-interest rate

²⁵This is also pointed out by Levin and Davidson (2005), who note that “[a] single-dimensional risk analysis would allow for hedging prepayment risk by combining premium MBS and discount MBS, a strategy any experienced trader knows would fail.”
sources of prepayment risk. The multiplier $\Lambda$ is also allowed to vary over time. Because of the large number of parameters that characterize the $P$ prepayment function, we make the assumption that the $Q$ prepayment function is a multiple of the $P$ prepayment function for parsimony. As we discuss below, our key finding that the smile pattern in OAS is due to prepayment risk premia while $OAS^Q$ is flat with respect to moneyness is preserved in an alternative framework where the $Q$ prepayment is estimated from the cross section of MBS prices only.

Pricing information on a standard pass-through MBS alone is insufficient to identify $\Lambda$, because a single observable (the price) can only determine one unknown (the spread) in the pricing model, leaving $\Lambda$ unidentified. To resolve this identification problem, we use additional pricing information from “stripped” MBS, which separate cash flows from pass-through securities into an interest component (“interest only” or IO strip) and a principal component (“principal only” or PO strip). Cash flows of these strips depend on the same underlying prepayment path and therefore face the same prepayment risk, but are exposed to it in opposite ways, as illustrated in Figure 4. As prepayment rates increase (top to bottom panel), total interest payments shrink (since interest payments accrue only as long as the principal is outstanding) and thus the value of the IO strip declines. Conversely, principal cash flows (the gray areas) are received sooner, and therefore the value of the PO strip increases.

We exploit the differential exposure of the two strips to prepayments to identify $OAS^Q$ and $\Lambda$, as illustrated graphically in the example of Figure 5. At $\Lambda = 1$, the physical prepayment speed, the OAS on the IO strip (shown in black) is about 200 basis points and the OAS on the PO (shown in gray) is about zero. As $\Lambda$ increases, the OAS on the IO declines while the spread on the PO increases because of their opposite sensitivities to prepayments. The sensitivity of the OAS on the pass-through (“recombined” as the sum of IO and PO) is also negative (red line), because in this example it is assumed to be a premium security and so its price declines with faster prepayments.

For each IO/PO pair, we identify $\Lambda$ as the crossing of the OAS IO and PO schedules at the point where the residual risk premium ($OAS^Q$) on the two strips is equalized.\textsuperscript{26} By the law of one price, the residual risk premium on the pass-through will also be equalized at this point; thus, the OAS schedule on the pass-through intersects the other two schedules at the same point. The difference between the OAS on the pass-through at the physical prepayment speed ($OAS$) and at the market-implied one ($OAS^Q$) is then equal to the prepayment risk premium paid on the

\textsuperscript{26}MBS market participants sometimes calculate “break-even multiples” similar to our $\Lambda$ but, to our knowledge, do not seem to track them systematically as measures of risk prices.
pass-through. More formally:

**Proposition 4.1.** Consider a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\), generated by an \(N\)-dimensional Brownian motion \(Z_t\), that generate interest rate risk and non-interest-rate prepayment risk, and a Poisson process \(J_t\), that generates liquidity risk for the market. Assume that all securities written on pool \(j\) lose the same fraction \(\alpha_{jt}\) of their market value in the case of a jump occurring at date \(t\), so that the IO strip and the PO strip on a pool \(j\) have equal exposures to non-prepayment sources of risk. Then, by no arbitrage, when expectations are calculated under the measure risk-neutral with respect to all Brownian shocks, \(Q_{r,\gamma}\), the remaining risk premia are equalized on the strips and recombined pass-through, so that

\[
OAS_{Q,\text{IO}_j} = OAS_{Q,\text{PO}_j} = OAS_{Q,j}.
\]

The \(OAS_{Q,j}\) on security of type \(k \in \{\text{IO}, \text{PO}, \text{pass-through}\}\) is defined implicitly by

\[
P_{k,jt} = \mathbb{E}^{Q_{r,\gamma}}_{t} \left[ \int_{0}^{\infty} \exp \left( -\int_{0}^{s} \left[ r_{t+u} + OAS_{Q,k,jt+u} \right] du \right) \, dX_{k,jt+s} \right],
\]

where \(P_{k,jt}\) is the market price of the security, \(\{dX_{k,jt+s}\}\) is the stream of cash-flows earned by the security, and \(r_t\) is the risk-free rate, which evolves as a function of \(Z_t\).

**Proof.** See the derivation of the \(OAS_{Q}\) in Appendix E.

We then define the prepayment risk premium component in the OAS as follows:

**Definition 4.1.** The prepayment risk premium on a pass-through security (consisting of the combination of an IO and PO strip on the same underlying pool) is equal to \(OAS \sim OAS_{Q}\).

We apply this method to each IO/PO pair in our sample, thereby identifying pool- and date-specific \(\Lambda\) and \(OAS_{Q}\). This allows us to study time-series and cross-sectional variation in the \(OAS_{Q}\) without imposing parametric assumptions and we can thus remain agnostic as to whether prepayment risk or other risks are the source of the OAS smile. Notice that while the above discussion (and the theoretical derivations in Appendix E) has emphasized the case of prepayment risk premia due to non-interest rate factors, our empirical setting also subsumes the possibility of a prepayment event risk premium (which by itself would lead to date-specific \(\Lambda\)).

27 An alternative way to identify \(\Lambda\) would be to assume that the OAS reflects only prepayment risk. With this approach, Levin and Davidson (2005) obtain a \(Q\) prepayment function by equalizing the OAS (relative to agency debt) on all pass-through coupons to zero. By construction, both the time-series and cross-sectional variation in the OAS will then be the result of variation in prepayment risk.
The key to this identification is the assumption that $OAS^Q$ are equal across IO and PO strips on the same pool. One could relax this assumption by imposing a parametric form linking $OAS^Q$ (or $\Lambda$) across pools. That said, the impact on the prepayment risk premium and $OAS^Q$ on the pass-through will be limited for reasonable liquidity differences between IOs and POs. For example, we find that assuming $OAS^Q_{PO}$ to be 50 basis points higher than $OAS^Q_{IO}$ never changes $OAS^Q$ by more than 5 basis points relative to the baseline specification with $OAS^Q_{IO} = OAS^Q_{PO}$. Intuitively, as shown in Figure 5, the slope of the $OAS$ schedule for the pass-through is less steep in $\Lambda$ than the slopes of the IO and PO schedules, and thus $OAS^Q_{IO} - OAS^Q_{PO}$ differences will have a limited effect on the recombined pass-through.

### 4.2 Stripped MBS data

To implement the identification described above, we start with an unbalanced panel of end-of-day price quotes on all IO/PO pairs (“trusts”) issued by Fannie Mae, obtained from a large dealer, for the period January 1995 to December 2010.\(^{28}\) We merge these with characteristics of the underlying pools, using monthly factor tape data describing pool-level characteristics obtained from the data provider eMBS. We use end-of-month prices, which we also subject to a variety of screens, as described in Appendix B. Following these data-quality filters, our data include 3713 trust-month observations, or about 19 per month on average, from 95 trusts total. The year with the lowest number of observations is 1999, where we have an average of 10 trusts per month, while after 2005 we have at least 20 trusts in all but one month.

The original face value of securities in our sample ranges from $200 million to about $4.5 billion, with a median of $2 billion. The median remaining principal balance of trusts in months in our dataset is $1.13 billion. In the cross-sectional analysis, we average spread measures to the coupon level (weighting by market value of the trusts), resulting in 1005 coupon-month pairs that cover most of the outstanding coupons in the Fannie Mae fixed-rate MBS universe (on average, 91 percent of remaining face value).\(^{29}\) A potential concern is that the IO/PO strips we have are not

\(^{28}\)We end our sample on that date because, according to market participants, IO/PO strips became less liquid after 2010, as trading started focusing on Markit’s synthetic total return swap agency indices IOS, POS and MBX instead. These indices mimic the cash flows of strips on a certain coupon-vintage (e.g. Fannie Mae 30-years with coupon 4.5 percent originated in 2009). The methods in this paper could easily be extended to those indices.

\(^{29}\)As in Figure 1, this means that the range of trust coupons in which the remaining face value is concentrated shifts downward over time. For instance, in January 1995, about 90 percent of the face value of securities for which we have quotes is in 7, 7.5, or 8 percent coupon securities. In January 2003, over 90 percent are in 5.5, 6, 6.5, or 7 percent securities. Finally, in December 2010, the last month in our data, the most prominent coupons are 4, 4.5, 5, and 5.5, which together account for 88 percent of face value.
necessarily representative of securities traded in the TBA market, to which we are comparing our model output. As we will see, however, we obtain similar spread patterns based on IO/PO prices, both in the time series and cross section. One advantage of the stripped MBS that we are using relative to TBAs, which trade on a forward “cheapest-to-deliver” basis, is that we do not need to make assumptions about the characteristics of the security.

### 4.3 Interest rate and prepayment model

A standard MBS pricing model has two main components: an interest rate model and a prepayment model. The two are combined to simulate interest rate paths and corresponding prepayment flows to obtain model prices and spreads. We use a three-factor Heath et al. (1992) interest rate model, calibrated at month-end to the term structure of swap rates and the interest rate volatility surface implied by the swaption matrix, by minimizing the squared distance between the model-implied and the observed volatility surface. We obtain swap zero rates from an estimated Nelson-Siegel-Svensson curve. Details on the interest and yield curve model are provided in Appendix F.\(^{30}\)

The academic literature has considered either structural/rational prepayment models (e.g., Dunn and McConnell, 1981a,b; Stanton, 1995) or reduced-form statistical prepayment models estimated on historical data (e.g., Richard and Roll, 1989; Schwartz and Torous, 1989). While structural models are more appealing, MBS investors favor reduced-form models (see, e.g., Fabozzi, 2016), for example, because in tranched CMOs, cash flows depend on prior prepayments, whereas structural models are solved by backward induction (McConnell and Buser, 2011). We follow standard industry practice and use a reduced-form prepayment model.

The exact details of practitioner models are not publicly available, but they vary in the choice of controls and weighting rules for historical data, and often make ad-hoc adjustments to incorporate likely effects of expected or announced policy changes affecting prepayments (for instance, the Home Affordable Refinance Program in 2009 or the introduction of additional agency fees on new mortgages since 2007). Therefore, in order to better capture market participants’ expectations and be consistent with their pricing and spreads, we do not estimate our model on historical data, but

---

\(^{30}\)A potential concern with using the risk-neutral evolution of interest rates inferred from the swaption matrix for pricing MBS is limits-to-arbitrage between the interest rate swap and MBS markets. Such differences could translate into differential OAS\(^Q\) across moneyness. Since we do not impose cross-moneyness restrictions on OAS\(^S\), our empirical specification is sufficiently flexible to capture these effects. As we will see, however, the cross-section of OAS\(^S\) is flat in moneyness, both in the full sample and in the pre-crisis period, suggesting that swaptions and MBS are priced fairly relative to each other.
instead extract prepayment model parameters from a survey of dealer models from Bloomberg LP. In these surveys, major MBS dealers provide their model forecasts of long-term prepayment speeds under different constant interest rate scenarios (with a range of +/- 300 basis points relative to current rates). Carlin et al. (2014) use these data to study the pricing effects of investors’ disagreement measured from “raw” long-run prepayment projections. We, instead, extract model parameters of a monthly prepayment function by explicitly accounting for loan amortization, the path of interest rates, and changes in a pool’s borrower composition.

Prepayment sensitivities to interest rates and other factors differ over time and across securities, and we thus estimate model parameters specific to each security and date. We model the date \( \tau \) single-month mortality rate (SMM), which is the fraction of a pool that preps, of security \( j \) to match the average projected long-run survey speed for the different interest rate scenarios. These scenarios provide information on a pool’s prepayment sensitivity to the incentive to refinance (INC\( _{\tau} \)). The functional form of our prepayment model is:

\[
s^{j}_{\tau} = \chi^{j}_{\tau}s^{1}_{\tau} + \left(1 - \chi^{j}_{\tau}\right)s^{2}_{\tau}, \text{ for } t < \tau \leq t + \text{Maturity}_j
\]

where

\[
s^{i}_{\tau, \tau} = b^{i}_{1}\min \left(\text{Age}^{j}_{\tau}/30, 1\right) + \kappa_{i} \cdot \frac{\exp \left(b^{i}_{2} + b^{i}_{3} \cdot \text{INC}^{j}_{\tau}\right)}{1 + \exp \left(b^{i}_{2} + b^{i}_{3} \cdot \text{INC}^{j}_{\tau}\right)} \text{ for } i = 1, 2.
\]

This functional form allows us to capture a key feature of the time evolution of MBS prepayments: the so-called burnout effect, which is the result of within-pool heterogeneity in borrowers’ sensitivity to the refinancing incentive. Because more sensitive borrowers are the first to exit the pool when rates decline, the pool’s overall sensitivity to interest rates drops over time even if interest rates are unchanged. To capture this effect, we assume the pool is composed of two types of borrowers: fast refinancers (group 1) and slow refinancers (group 2), with respective shares \( \chi_{\tau} \) and \( 1 - \chi_{\tau} \) and shares \( s^{1}_{\tau} \) and \( s^{2}_{\tau} \). This setup is a simplified version of the heterogeneous refi-

---

31 Until May 2003, dealers provided a single set of forecasts for each coupon (separately for Fannie Mae, Freddie Mac, and Ginnie Mae pass-through securities); since then, they provide separate forecasts for different vintages (for instance, a 5.5 percent coupon with average loan origination date in 2002 versus a 5.5 percent coupon with origination in 2005).

32 A notable detail is that in our model, we define INC as the end-of-month 10-year swap rate minus the pool’s weighted average coupon (WAC). This is different from the “true” interest rate incentive faced by a borrower, which would be the mortgage rate minus WAC. However, our formulation has the major advantage that it does not require us to specify a model for the gap between mortgage rate and swap rate. The average gap between 30-year FRM rate and the 10-year swap rate over our sample period is about 1.2%.

33 In the extreme case, some borrowers never refinance even when their option is substantially in the money. Possible reasons for this non-exercise of the prepayment option include unemployment or other credit problems (Longstaff, 2005) or a lack of financial sophistication (sometimes called “woodhead” behavior; Deng and Quigley, 2012).
nancing cost framework of Stanton (1995). As shown in equation (4.1), total pool prepayments are share-weighted averages of each group’s prepayment speed. Each group’s prepayment depends on two components. The first, which is identical to both groups, is governed by \( b_1 \) and accounts for non-rate-driven prepayments, such as housing turnover. Because relocations are less likely to occur for new loans, we assume a seasoning of this effect using the industry-standard “PSA” assumption, which posits that prepayments increase for the first 30 months in the life of a security and are constant thereafter. The second component captures the rate-driven prepayments due to refinancing. This is modeled as a logistic function of the rate incentive (INC), with a sensitivity \( \kappa_i \) that differs across the two groups: \( \kappa_1 > \kappa_2 \). Since group 1 prepays faster, \( \chi \tau \) declines over time in the pool. This changing composition, which we track in the estimation, captures the burnout effect. We provide more detail on the prepayment model and parameter estimation in Appendix F.

Figure 6 shows estimated prepayment functions for different loan pool compositions and using average parameters \( b_1, b_2, b_3 \) across all securities in our sample. Prepayments (at an annual rate, known as the “constant prepayment rate,” or CPR) display the standard S-shaped prepayment pattern of practitioner models. They are not very sensitive to changes in interest rates (and thus INC) for securities that are deeply ITM or OTM, but highly sensitive at intermediate moneyness ranges. The black (top) line shows that a pool with \( \chi = 1 \) reaches a maximum predicted CPR of about 75 percent when it is deeply ITM, in contrast to only 35 percent when the share of fast refinancers is only 0.25 (red line). Thus, the changing borrower composition, even with a constant INC, implies a decline in prepayments over time because of the pool’s burnout (decreasing \( \chi \)).

5 Model results

The pricing model produces the standard OAS measure as well as the \( \text{OAS}^Q \), which is adjusted for (or risk-neutral with respect to) not only interest rate risk but also prepayment risk. In this section we present the output of the model in terms of spreads in the cross section and time series. We then relate average \( \text{OAS}^Q \) and prepayment risk premia to fixed-income and MBS-specific risk measures in order to help interpret model results and variation in MBS spreads. We finally discuss the response of MBS spreads to the Fed’s first LSAP announcement in November 2008.
5.1 OAS smile

The cross-sectional results are summarized in Figure 7. Similar to our findings for the TBA spreads (Figure 2), the OAS from our model exhibits a smile in moneyness (panel a): spreads are lowest for securities with moneyness near zero and increase for securities that are either OTM, or especially, ITM. As shown in panel (b), the $OAS_Q$, which strips prepayment risk from the OAS, does not appear to vary significantly with moneyness, suggesting that differences in liquidity do not contribute to the OAS smile. Instead, as shown in panel (c), the difference between the OAS and $OAS_Q$ closely matches the smile pattern in the OAS; in other words, the differential exposure to prepayment risk explains the cross-sectional pattern in the OAS. Additionally, panel (d) displays the difference in implied long-run prepayment speeds between the risk-neutral (Q) and physical (P) prepayment models. OTM securities tend to have slower risk-neutral speeds, while ITM securities tend to have faster risk-neutral speeds. Thus, in both cases the risk-neutral model tilts the prepayment speeds in the undesirable direction from the point of view of the investor. That is, market prices imply that prepayments are faster (slower) for securities that suffer (benefit) from faster prepayments, which is exactly what one would expect as market-implied prepayments include compensation for risk.

In Table 4, we use regressions to study if these cross-sectional patterns in the two components of OAS are robust to including month fixed effects (in order to focus on purely cross-sectional variation) and to ending the sample before the financial crisis period by sorting the different coupons in bins by moneyness, as in the earlier Table 2. Panel (a) of Table 4 shows that there is little systematic pattern in $OAS_Q$ across bins; results in columns (2) and (4) suggest in fact that ATM coupons may have slightly higher $OAS_Q$ than the surrounding coupons, but the differences are small. There is some evidence that the most deeply ITM coupons (moneyness $\geq$ 2.25) may command a positive premium, which could be driven by the reduced liquidity of these (generally very seasoned) coupons. Turning to the prepayment risk premium, panel (b) of Table 4 shows that the (slightly tilted) smile pattern shown in panel (c) of Figure 7 is robust to adding month fixed effects and excluding the financial crisis period. The coefficients suggest that the magnitude of the prepayment risk premium is economically meaningful: securities that are 1.25 percentage points or more ITM command a premium of 20 basis points (annual) or more relative to ATM securities.

In sum, while the prepayment risk premium in the cross section is strongly linked to the mon-

---

34 We use fewer bins because our IO/PO strips have less coverage of very deeply OTM (moneyness $<-1.75$) or ITM (moneyness $>2.75$) coupons.
eyness of the securities, we find little evidence that this is also the case for the remaining risk premium \( OAS^Q \), suggesting that differential liquidity across coupons is likely not a major driver of cross-sectional variation in spreads (except perhaps for the most deeply ITM securities). Instead, the smile in prepayment risk premia suggests that both ITM and OTM securities earn positive compensation for prepayment risk, consistent with them not being hedges for one another. Going back to our conceptual framework, this would arise if either a single shock drives prepayments but with opposite effects on ITM and OTM securities (case 2), or, perhaps more realistically, if OTM and ITM securities were subject to distinct but independent shocks (case 3). For instance, OTM prepayments could be primarily driven by housing-relocation shocks, such as house prices, whereas variation in ITM prepayments could be due to shocks to refinancing activity.

Though Table 4 shows that the cross-sectional patterns in the two components of OAS are robust to including month fixed effects and to ending the sample in 2007, a potential concern is whether these results could be due to potential misspecification of the risk-neutral prepayment function as a multiple of the physical prepayment function. Note that the multiple \( \Lambda \) is security and date specific, so that our estimate of the \( Q \) prepayment function is semi-parametric and allows us to capture risk premia with the least restrictive prior assumptions. Nevertheless, in Appendix G we explore an alternative identification methodology that estimates the risk-neutral prepayment function directly from IO/PO prices only, without requiring us to start with a \( P \) prepayment function. While this alternative approach does not allow the estimation of prepayment risk premia, it confirms this section’s finding that the \( OAS^Q \) does not vary systematically with moneyness.

### 5.2 Time-series variation

We now turn to the variation in the average OAS in the time series. As in Section 2, we construct a market value-weighted index of our model-implied OAS.\(^{35}\) Comparing the OAS in Figure 8 to the corresponding series in panel (a) of Figure 2 confirms that our model output is close to its dealer counterparts. As in these models, the level of the average OAS is generally close to zero, but increases in periods of market stress. Further, our pricing model finds the difference between OAS and \( OAS^Q \) to be small and the two series to tightly co-move, meaning that much of the OAS variation results from changes in \( OAS^Q \) (gray line). Thus, although it is an important determinant of the cross-sectional variation in spreads, prepayment risk does not appear to be the dominant

\(^{35}\)We do this by first averaging spreads across trusts within a coupon (weighting each trust by its market value, given by its remaining principal balance times the sum of the prices of its IO and PO strips). Then, we average across coupons in a given month (weighting each coupon by its market value based on TBA prices, as in Section 2.2).
driver of the OAS time-series variation. Indeed as shown in Figure 1, the share of deeply OTM or ITM securities, which earn most compensation for prepayment risk, is limited; this arises because most securities are close to ATM when issued.\textsuperscript{36} However, prepayment risk in the MBS universe can be significant when mortgage rates move sharply, as in early 1998, the summer of 2003, and in 2009 and 2010 as mortgage rates reached historic lows and the gap between the average OAS and $OAS^Q$ widened.

We next investigate the determinants of the time-series variation in OAS, and in particular its two components: $OAS^Q$ and the prepayment risk premium. Table 5 shows results from monthly regressions of the OAS, and its components, on mortgage-specific risk factors, such as spreads on agency debt (or debentures) relative to swaps, agency MBS issuance (normalized by broker-dealer book equity, and subtracting Fed MBS purchases in 2009 and 2010), the average squared moneyness of the MBS universe, as well as dealer disagreement about future prepayment speeds, which Carlin et al. (2014) find to be significant predictors of MBS returns.\textsuperscript{37} We construct our disagreement measure in the same way as Carlin et al. (2014), but based on the Bloomberg surveys on Fannie Mae prepayment speeds that we already use for the physical prepayment model (while Carlin et al. instead use forecasts of Ginnie Mae prepayments). We also include the credit spread, which was the main economically and statistically significant factor in the TBA analysis in Table 1.

We find that average $OAS^Q$ are related to spreads on (unsecured) agency debentures. As noted earlier, agency MBS are typically perceived as being free of credit risk, but since the government guarantee on securities issued by Fannie Mae is only implicit, investors’ perceptions of this guarantee (along with the perceived credit risk of agencies) may change over time and thus affect both spreads on agency debt and MBS. In particular, both $OAS^Q$ and agency debt spreads increased in the fall of 2008, when Fannie Mae and Freddie Mac were placed in conservatorship by the U.S. Treasury. The spreads on MBS and agency debt do, however, also co-move earlier in the sample, pointing to other common factors such as liquidity and funding costs of these securities.\textsuperscript{38}

\textsuperscript{36}After issuance, the moneyness of securities fluctuates as a function of interest rates and the remaining balance declines with prepayments, lowering the importance of older issues on a value-weighted basis.

\textsuperscript{37}Agency debenture yields and swap yields come from Barclays; agency MBS issuance is from eMBS; Fed MBS purchases from the \url{https://www.federalreserve.gov/regreform/reform-mbs.htm}; and broker-dealer equity from Table L.130 in the Flow of Funds (interpolated linearly between quarter-ends).

\textsuperscript{38}The spread between Fannie Mae debentures and Treasury bonds of equal maturity fell following the conservatorship announcement, but then substantially increased through the end of 2008. Since there should have been essentially no difference in the strength of the debt guarantee between debentures and Treasuries at that point, and since the spread widening was stronger for shorter maturity bonds, Krishnamurthy (2010) argues that this reflects a flight to liquidity. In line with this interpretation, our $OAS^Q$ also reaches substantially higher levels in October compared to August 2008, despite the reduction in credit risk to investors.
Credit spreads (Baa-Aaa) continue to be significantly related to OAS, mostly through $OAS^Q$ rather than the prepayment risk component. The sensitivity of $OAS^Q$ to credit spreads suggests common pricing factors in the MBS and credit markets, such as limited risk-bearing capacity of financial intermediaries (see, for example, Shleifer and Vishny, 1997; Duffie, 2010; Gabaix et al., 2007; He and Krishnamurthy, 2013). In these models, financial intermediaries are marginal investors in risky assets; when their financial constraints bind, their effective risk aversion increases, raising risk premia in all markets. Thus, when the supply of risky assets relative to intermediaries’ capital decreases, financial constraints are relaxed, lowering required risk compensation. In line with these predictions, we find that higher supply of MBS, measured by issuance relative to mark-to-market equity of brokers and dealers, also positively correlates with average $OAS^Q$.\footnote{As noted above, we net out monthly Fed purchases by settlement month over 2009-10 from new issuance to more properly measure fluctuations in net supply that is absorbed by investors. Relatedly, GSEs’ conservatorship agreements have required them to divest their portfolios since 2010, a process that would affect net supply in the hands of other investors. However, during the first round of LSAPs, which we focus on, changes in GSE holdings of agency MBS were quite small relative to Fed purchases and issuance.} We explore this channel further in the next section, where we study the effects of Fed MBS purchases, which absorb supply in the hands of investors, on the OAS and its components.

Disagreement about future prepayments (for given rates) is positively related to the prepayment risk premium, in line with the findings of Carlin et al. (2014); however, the economic magnitude of the coefficient is relatively small. Finally, as previously discussed, the OAS smile implies that spreads, and in particular their prepayment risk component, are largest for deeply OTM and ITM securities. This suggests that when the market-value weighted moneyness is either very positive or very negative, the average OAS and prepayment risk premium should be large. In line with this prediction, we find average squared moneyness to be positively related to the average prepayment risk premium.

5.3 Interpreting the OAS response to the Fed’s LSAPs

As discussed above, MBS spreads are positively related to MBS supply, a finding that is consistent with intermediary asset pricing models with limited risk-bearing capacity. In this section we provide additional evidence on this channel by focusing on the Fed’s large-scale asset purchase (LSAP) program. The program has entailed an unprecedented shift in the composition of the MBS investor base as the Fed now holds more than a quarter of the total agency MBS universe—up from nothing prior to the financial crisis. We decompose spreads using our pricing model and show how our model can explain the divergence in OAS across different coupons following the
initial announcement of the program.

We focus on spread movements after November 25, 2008, when the Fed announced its first round of purchases of up to $500 billion in agency MBS. Based on the current coupon MBS, which is the focus of much of the research on this topic—with the important exception of Krishnamurthy and Vissing-Jorgensen (2013) which we discuss in the paper’s conclusion—the announcement had a substantial effect on the MBS market (see, e.g., Gagnon et al. 2011 or Hancock and Passmore 2011; Stroebel and Taylor 2012 are more skeptical). According to different dealer models, the current coupon OAS, which had been at record levels of 75–100 basis points over October and November 2008, fell 30–40 basis points on the day of the announcement, and stayed around the lower level afterwards. Consistent with the decline in secondary MBS spreads and yields, headline 30-year FRM rates dropped nearly a full percentage point between mid-November and year-end 2008.

Spread movements on the current coupon MBS alone hide significant heterogeneity across the coupon stack, as evidenced by the series in Figure 9, which are median spreads across dealer models (the same used in Section 2) for the four main coupons traded at that time. OAS that were all at similarly elevated levels in the fall of 2008 diverged following the announcement. Between October and November 2008, spreads on low coupons (4.5 and 5) fell, while, over the same period, those on higher coupons were little changed and then even widened through the end of December. Since high coupons represent the majority of outstanding MBS, this implies that, for specialized investors in this market, the recapitalization effect of monetary policy described in Brunnermeier and Sannikov (2012) was limited.

The earlier findings from our model suggest two potentially countervailing effects of Fed MBS purchases on OAS. On the one hand, Fed purchases reduce MBS supply to be absorbed by risk-sensitive investors, thereby reducing the required risk premium on all MBS (through $OAS^Q$). On the other hand, movements in mortgage rates associated with such purchases alter securities’ moneynesses, shifting the OAS along the smile and changing the prepayment risk premium.

The results from our model are shown in the bottom panel of Figure 10. First, OAS movements (in black) for IO/PO pass-throughs are similar to the TBA ones. In terms of the MBS supply

---

40 The Fed also announced that it would purchase up to $100 billion in agency debt. The purchases began in early January 2009. The program was then extended in March 2009, when it was announced that an additional $750bn in agency MBS, $100bn in agency debt, and $300bn in long-term Treasuries would be purchased over the following year.

41 Over several months in early 2009, monthly Fed MBS purchases were absorbing essentially all new agency MBS issuance.

42 The strips we have available do not necessarily have the same characteristics as what the dealers assume to be cheapest-to-deliver in TBA trades; therefore, our OAS levels do not exactly line up with theirs for all coupons in all months. Nevertheless, patterns are very similar, especially in changes.
effect, we discussed above how the $OAS_Q$ component is flat across coupons and declines with a reduction in MBS supply. Consistent with this, we find that the $OAS_Q$ evolves similarly for the 4.5, 5, and 5.5 coupons. For the 6 coupon, $OAS_Q$ increases in November and December, before dropping toward the level of the other coupons in January as actual LSAP purchases begin. The $OAS_Q$ effect thus suggests that the LSAP program lowered non-prepayment risk premia across the coupon stack.\footnote{In addition to the supply effect, the Fed announcement may also have strengthened the perceived government backing of Fannie Mae and Freddie Mac and improved the liquidity of agency securities (Hancock and Passmore, 2011; Stroebel and Taylor, 2012).}

The cross-coupon “homogeneous” $OAS_Q$ impact of the Fed’s policy is, however, masked by changes in the prepayment risk premia that vary with MBS moneyness, shown in the top panel of Figure 10. The 4.5 starts out OTM and moves ATM as mortgage rates drop, while the 5.5 and 6, which are around ATM in October, move quite deeply ITM. Based on the OAS smile, the 4.5 should command a prepayment risk premium prior to November and the 5.5 and 6 coupons afterward. The bottom panel shows that this is indeed the case: the narrowing in the gap between the black and gray lines means that the decrease in the OAS of the 4.5 coupon is in part due to the decrease in its prepayment risk exposure following the drop in rates. In contrast, the prepayment risk premia on the 5.5 and 6 coupons are high from December onward as they move deeply ITM and are more sensitive to prepayment risk.

In sum, increases in the moneyness of high coupons following the November 2008 LSAP announcement led to an increase in their prepayment risk premium, which explains why their OAS did not fall, even though $OAS_Q$ declines across the coupon stack as the Fed started absorbing MBS supply. We discuss the policy implications of these results below.

6 Conclusions

Our pricing model has two main implications. In the cross section, risk premia associated with non-interest-rate prepayment factors explain the OAS smile, which is the fact that OAS tend to be lower for at-the-money MBS than for others. In the time series, the model implies that the average OAS is primarily driven by non-prepayment risk factors linked to credit spreads, MBS supply, and spreads on other agency debt. These results suggest that risk-bearing capacity of MBS investors, and the liquidity and default risk of agency securities, drive aggregate spread variation and are important determinants of homebuyers’ funding costs.
A number of studies have suggested that the Fed’s asset composition can influence financial asset prices as postulated in Bernanke (2009). Gagnon et al. (2011), building on Greenwood and Vayanos (2014), argue that Fed purchases of any long-term asset (that is, either MBS or Treasuries) affect term premia on fixed income assets by reducing the market price of duration risk. On the contrary, Krishnamurthy and Vissing-Jorgensen (2011) argue that Fed purchases have more nuanced pricing implications and mainly affect the price of the securities being purchased, rather than operating through a more general effect on duration risk premia.

The OAS declines following the first LSAP announcement in November 2008 are consistent with Fed purchases having a disproportionate effect on the targeted assets as opposed to operating through a pure monetary or signaling channel. Moreover, OAS on low coupons fell more than those on high coupons. Our model attributes these facts to: (i) the OAS falling (roughly) equally across coupons as the Fed absorbs supply and lowers the risk premia required by specialized investors; and (ii) high coupons moving deeper ITM, which increases the prepayment risk premium on those coupons and prevents their total OAS from falling.

Differential responses across MBS coupons are not specific to monetary policy changes in 2008/9, a period in which severe market disruptions may have affected the reactions to Fed announcements. For example, Krishnamurthy and Vissing-Jorgensen (2013) discuss the “taper tantrum” around the June 19, 2013 FOMC meeting, when rates backed up on investors’ fears that the Fed would start reducing its purchases soon. Around this event, the OAS increased substantially for low coupons, while the OAS on higher coupons stayed almost unchanged. Krishnamurthy and Vissing-Jorgensen interpret this latter fact as evidence that capital constraints (or limited risk-bearing capacity) are unimportant at that time. Instead, they argue that the increase in lower-coupon OAS comes about because the “scarcity effect” for low coupons weakens as the anticipated Fed demand for those coupons decreases. Based on this interpretation, they suggest that the Fed could sell the higher coupons from its MBS portfolio, should it target a lower balance sheet size, without causing an increase in production-coupon OAS.

Our model, which does not rely on cross-coupon segmentation, suggests a different explana-

\footnote{In contrast, they argue that the “capital constraints” channel was the main channel responsible for the decrease in OAS following the November 2008 announcement, in line with our discussion in the previous section.}

\footnote{Krishnamurthy and Vissing-Jorgensen’s MBS scarcity channel, which is specific to the TBA market, implies larger price responses for coupons directly targeted by Fed purchases. According to this channel, as demand for a specific coupon increases, the quality of pools delivered in the TBA contract (as measured by their prepayment characteristics) improves, so that the equilibrium price increases to elicit pool delivery. Because this scarcity channel works at the level of each coupon, it predicts that Fed purchases do not affect risk premia on non-targeted MBS, such as higher coupon TBAs or MBS not deliverable in the TBA market.}
tion for the reaction to the June 2013 events: the increase in the quantity of securities that non-Fed investors have to hold because of the anticipated taper increases the required risk premium (through $OAS^Q$) on all MBS; however, because rates increase at that point, the prepayment risk premium on high coupons (that were previously deeply ITM) falls as they become ATM, so that their overall OAS remains roughly constant. Because of the differential prepayment risk exposure across MBS, the stability in high-coupon OAS around this event is thus not evidence of a lack of capital constraints for MBS investors, implying that potential sales of high coupons might still increase OAS on lower coupons and therefore increase mortgage rates.

From a broader perspective, this paper provides evidence for intermediary asset pricing in fixed income markets. Recent literature (such as He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014) has proposed that intermediaries’ risk-bearing capacity impacts risk premia during periods of market stress. While much of the previous discussion has focused on the response of the $OAS^Q$ to Fed purchases in the wake of the crisis, the $OAS^Q$ reacts to changes in the outstanding supply of MBS and credit spreads even during normal market conditions. The latter finding is consistent with theories (e.g. Gromb and Vayanos, 2002; Brunnermeier and Pedersen, 2009; Adrian and Boyarchenko, 2012) that link risk premia to intermediary balance sheet constraints even in periods when intermediaries are well capitalized.

Finally, our work contributes to the literature studying the causes and consequences of mortgage market design (Campbell, 2013). The key distinguishing feature of the U.S. mortgage market is the pre-dominance of the 30-year fixed rate mortgage contract without prepayment penalties, which heavily relies on the agency securitization market that we studied in this paper (Green and Wachter, 2005; Fuster and Vickery, 2015). Absence of prepayment penalties does of course not imply a costless prepayment option for borrowers, as this optionality is priced ex-ante by investors. Indeed, the prepayment risk premium in the OAS analyzed in this paper is a direct manifestation of this cost. More generally, the absence of prepayment penalties also impacts mortgage rates through other equilibrium channels such as higher term premia and rate volatility (Malkhozov et al., 2016; Hanson, 2014). Evaluating the importance of all these factors remains an open research question, which is central to understanding the costs and benefits of different mortgage contract designs in the U.S. and abroad.

---

46The 10-year Treasury yield increased by 40 basis points from June 18 to June 25; the Freddie Mac headline FRM rate even increased by more than 50 basis points.
References


Figure 1: Share of total MBS remaining balance by coupon against the current coupon rate. The shaded gray areas are the share (left axis) of the total remaining principal balance (RPB) in each 30-year fixed-rate Fannie Mae MBS coupon relative to the total RPB of all coupons. RPBs are from eMBS. The red line is the current coupon (the interpolated coupon that trades at par) in the TBA market (right axis), obtained from dealer data.
Figure 2: Time-series and cross-sectional variation of the OAS based on dealer TBA data. The top panel displays the time series at monthly frequency of the option-adjusted spread (to swaps) on a value-weighted index based on TBA quotes from six dealers. The bottom panel displays a scatterplot and a local smoother of the cross-sectional variation in the OAS across MBS coupons as a function of their moneyness. Moneyness is calculated as the coupon rate plus 50 basis points (to account for servicing and the guarantee fee) minus the 30-year fixed-rate mortgage rate obtained from Freddie Mac. The figure only includes coupons with remaining principal balance of at least 100 million in 2009 dollars. All data is as of month-end and covers the period 1996-2010. Further details on the construction of the value-weighted series is reported in Section 2. Appendix B.1 discusses the sample in more detail.
Figure 3: Cross-sectional variation in hedged returns. The figure shows monthly $t \rightarrow t + 1$ hedged returns on MBS in the Bloomberg Barclays index by coupon against securities’ moneyness as of the end of month $t$. Moneyness is calculated as the coupon rate plus 50 basis points (to account for servicing and the guarantee fee) minus the 30-year fixed-rate mortgage rate obtained from Freddie Mac. Each dot represents mean returns for one of twenty equally sized moneyness bins. The gray line represents a local smoother fitted to the underlying data, that is, each coupon’s return. The sample period is January 1994 to June 2016.
Figure 4: MBS cash flows for different prepayment speeds. The colored areas represent (undiscounted) monthly cash flows for a hypothetical MBS with original principal of $100, loan rate of 4.5% and coupon of 4%. The 50 basis point difference between the loan and coupon rate is earned by the servicer and guaranteeing agency (blue area). Scheduled (amortization) and unscheduled (prepayment) principal payments are shown as gray areas. The sum of the two areas in each chart adds up to $100. As a result PO strips benefit from faster prepayments (early repayment). The IO strip receives the monthly interest payment (coupon rate at monthly rate) on the principal balance outstanding. The top (bottom) panel shows a slow (fast) constant prepayment rate scenarios. With fast prepayments, interest payments (red area) are much smaller, thus IOs suffer from fast prepayments. Calculations are based on the formulas in Appendix A.

(a) Slow prepayment (CPR = 12%)

(b) Fast prepayment (CPR = 24%)
Figure 5: Visualization of the identification assumption for $OAS^Q$. The figure shows the OAS on a IO, PO and pass-through for the same underlying collateral as a function of the prepayment multiple ($\Lambda$) on the physical prepayment speed. Higher $\Lambda$ increases the market-implied prepayment speed relative to the physical one. The OAS on the IO (PO) declines (increases) in $\Lambda$. This follows from the relation between the value of each strip and prepayments shown in Figure 4. The OAS on the pass-through in this example also declines in $\Lambda$ because the pass-through is a premium security ($p^{IO} + p^{PO} > 100$). The three OAS differ at the physical speed ($\Lambda = 1$) but are equalized at the risk-neutral speed. This value of $\Lambda$ defines the market implied prepayment speed and the $OAS^Q$. 
Figure 6: Survey-implied average prepayment rates at different levels of burnout. The figure shows the prepayment function in equation (4.2) with parameters $b_1$, $b_2$ and $b_3$ set at their sample means. Differences in $\chi$ parametrize the burnout effect. Fast refinancers (group 1) and slow refinancers (group 2) are present in the pool with shares $\chi$ and $1 - \chi$. As the share of more sensitive borrowers $\chi$ falls, the pool’s overall sensitivity to interest rates is reduced. The vertical axis is the “conditional prepayment rate” (CPR, or annualized prepayment rate) and the horizontal axis is a measure of the incentive to prepay the mortgage.
Figure 7: Cross-sectional variation in model-implied OAS and prepayment speeds. The panels show scatterplots and local smoothers of the cross-sectional variation in the model-implied OAS, OAS^Q, the prepayment risk premium (OAS – OAS^Q) and the difference between the market-implied and physical lifetime prepayment speed. All measures are for the recombined pass-throughs obtained as the sum of the value of the IO and PO components. Refer to Figure 5 for a summary of the relation between physical and market-implied spreads and prepayment speeds. The horizontal axis measures moneyness, which is calculated as the coupon rate plus 50 basis points (to account for servicing and the guarantee fee) minus the 30-year fixed-rate mortgage rate obtained from Freddie Mac. Additional details on the pricing model are given in Section 4.

(a) OAS

(b) OAS^Q

(c) OAS – OAS^Q

(d) Prepayment speeds: Q vs. P
Figure 8: Time series of the OAS on the pass-through index from IO/PO strips. This figure shows time-series variation in OAS and OAS$^Q$ on a value-weighted index computed from IO/PO prices. To construct the index, we first average spreads across trusts for each coupon using market weights. Then, for each month we average across coupons using each market value from TBA prices. For each trust, the value of the pass-through is the sum of the value of the IO and PO. Refer to Figure 5 for a summary of the relation between OAS and OAS$^Q$ and to Section 4 for more details on the pricing model.
Figure 9: Variation in OAS around the November 25, 2008 LSAP announcement by the Federal Reserve. This figure shows the evolution of the OAS for Fannie Mae MBS with coupons 4.5, 5.0, 5.5, and 6.0 percent, which were the most heavily traded at that time. OAS data is based on median TBA quotes from six dealers and reported as of month-end. The vertical line indicates November 25, 2008, which is the announcement date.
Figure 10: OAS decomposition around the November 25, 2008 LSAP announcement by the Federal Reserve. This figure shows the MBS moneyness by coupon (upper panels) and movements in OAS and OAS$^Q$ (bottom panels) based on IO/PO prices and our pricing model. Moneyness is calculated as coupon plus 50 basis points minus the 30-year fixed-rate mortgage rate (from Freddie Mac). Data is as of month-end. Vertical lines indicate November 25, 2008, when the Federal Reserve announced its large-scale asset purchase program.

(a) Moneyness across coupons

(b) Spreads across coupons
Table 1: Time-series regressions of OAS on the TBA index from dealer data. Coefficient estimates from OLS regression of the TBA index OAS (shown in the upper panel of Figure 2) on the Aaa-Treasury spread, the Baa-Aaa spread, the slope of the Treasury yield curve (difference between the 10-year and 3-month Treasury yield) and the 2-year into 10-year swaption implied volatility. The Aaa and Baa corporate bond yields are from Moody’s, while swaption volatility is from a dealer. Data is at monthly frequency and measured as of month-end. All regressors are standardized to have zero mean and unit standard deviation. Newey-West standard errors (6 lags) shown in brackets. Significance: * p < 0.1, ** p < 0.05, *** p < 0.01.

<table>
<thead>
<tr>
<th></th>
<th>(1) OAS</th>
<th></th>
<th>(2) OAS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa - Treas</td>
<td>3.2**</td>
<td>[1.5]</td>
<td>2.3*</td>
<td>[1.2]</td>
</tr>
<tr>
<td>Baa - Aaa</td>
<td>19.6***</td>
<td>[3.4]</td>
<td>9.7**</td>
<td>[3.8]</td>
</tr>
<tr>
<td>Treas Slope</td>
<td>-0.5</td>
<td>[2.4]</td>
<td>2.0</td>
<td>[1.3]</td>
</tr>
<tr>
<td>Swaption Vol.</td>
<td>2.7</td>
<td>[4.2]</td>
<td>-0.8</td>
<td>[2.4]</td>
</tr>
<tr>
<td>Const</td>
<td>8.8***</td>
<td>[1.8]</td>
<td>3.6***</td>
<td>[1.3]</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>0.76</td>
<td></td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>180</td>
<td></td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>Dates</td>
<td>199601.201012</td>
<td></td>
<td>199601.200707</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Cross section of OAS on TBA securities from dealer data by moneyness. Coefficient estimates from OLS regressions of the OAS (annualized, in basis points) on different moneyness bin dummies, either including or excluding calendar month fixed effects. Moneyness of a coupon $j$ at time $t$ is defined as Moneyness$_{j,t} = \text{Coupon}_j + 0.5 - \text{FRMrate}_t$. The dummy for the moneyness bin surrounding zero, $[-.25,.25)$, is the omitted category. Columns 1 and 2 in panel (a) show estimates for full sample 1996–2010; columns 3 and 4 exclude the period August 2007–December 2010 (thus excluding the financial crisis). In panel (b) the sample is split based on the moneyness of the MBS market: “Market ITM (OTM)” indicates that the balance-weighted average moneyness of all outstanding Fannie Mae MBS is $> 0 (< 0)$. In columns 1 and 2 of panel (b), the first two bins are merged since there are only 7 observations with moneyness $< -1.75$ when the market is ITM; similarly in columns 3 and 4 the last two bins are merged since there are only 5 observations with moneyness $\geq 2.75$ when the market is OTM. Robust standard errors (clustered at the month level) shown in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### (a) Full sample and pre-crisis sample

<table>
<thead>
<tr>
<th>Moneyness bin</th>
<th>Coefficient (1)</th>
<th>Coefficient (2)</th>
<th>Coefficient (3)</th>
<th>Coefficient (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full sample</td>
<td>Pre-crisis sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[-1.75, -1.25)$</td>
<td>8.9*** [1.9]</td>
<td>10.4*** [1.6]</td>
<td>7.7*** [1.2]</td>
<td>8.3*** [1.4]</td>
</tr>
<tr>
<td>$[-1.25, -.75)$</td>
<td>4.0*** [1.0]</td>
<td>5.3*** [1.1]</td>
<td>4.3*** [0.9]</td>
<td>4.8*** [0.9]</td>
</tr>
<tr>
<td>$.75, -2.5)$</td>
<td>-0.0 [0.4]</td>
<td>-0.1 [0.4]</td>
<td>0.8* [0.4]</td>
<td>0.8** [0.4]</td>
</tr>
<tr>
<td>$.75, 2.25)$</td>
<td>1.5*** [0.5]</td>
<td>1.5*** [0.5]</td>
<td>-0.1 [0.4]</td>
<td>-0.1 [0.4]</td>
</tr>
<tr>
<td>$.75, 2.75)$</td>
<td>6.4*** [1.2]</td>
<td>6.4*** [1.2]</td>
<td>2.6*** [0.9]</td>
<td>2.5*** [0.9]</td>
</tr>
<tr>
<td>$.75, 2.75)$</td>
<td>15.6*** [2.2]</td>
<td>15.0*** [2.1]</td>
<td>7.2*** [1.8]</td>
<td>7.2*** [1.7]</td>
</tr>
<tr>
<td>$.75, 2.75)$</td>
<td>27.2*** [3.7]</td>
<td>24.6*** [3.3]</td>
<td>11.9*** [3.0]</td>
<td>11.4*** [3.0]</td>
</tr>
<tr>
<td>$.75, 2.75)$</td>
<td>41.3*** [6.4]</td>
<td>33.9*** [5.5]</td>
<td>26.0*** [5.3]</td>
<td>23.2*** [5.2]</td>
</tr>
<tr>
<td>$\geq 2.75$</td>
<td>93.4*** [13.3]</td>
<td>85.4*** [10.9]</td>
<td>50.5*** [7.2]</td>
<td>47.1*** [7.2]</td>
</tr>
<tr>
<td>Const</td>
<td>7.5*** [1.5]</td>
<td>7.5 [0.8]</td>
<td>0.5 [0.8]</td>
<td></td>
</tr>
<tr>
<td>Month FE?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>0.22</td>
<td>0.61</td>
<td>0.24</td>
<td>0.38</td>
</tr>
<tr>
<td>Adj. R2 (within)</td>
<td>0.31</td>
<td>0.31</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>1532</td>
<td>1532</td>
<td>1151</td>
<td>1151</td>
</tr>
<tr>
<td>Dates</td>
<td>199601.201012</td>
<td>199601.201012</td>
<td>199601.200707</td>
<td>199601.200707</td>
</tr>
</tbody>
</table>

### (b) Subsamples based on moneyness of MBS market

<table>
<thead>
<tr>
<th>Moneyness bin</th>
<th>Market ITM</th>
<th>Market OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$&lt; -1.75$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$[-1.25, -.75)$</td>
<td>5.7*** [1.5]</td>
<td>7.9*** [1.6]</td>
</tr>
<tr>
<td>$.75, -2.5)$</td>
<td>0.3 [0.5]</td>
<td>0.2 [0.6]</td>
</tr>
<tr>
<td>$.75, 2.5)$</td>
<td>0.7 [0.7]</td>
<td>0.7 [0.7]</td>
</tr>
<tr>
<td>$.75, 2.75)$</td>
<td>8.9*** [2.5]</td>
<td>8.6*** [2.4]</td>
</tr>
<tr>
<td>$1.25, 1.75)$</td>
<td>17.4*** [3.6]</td>
<td>15.6*** [3.5]</td>
</tr>
<tr>
<td>$\geq 2.75$</td>
<td>98.5*** [13.4]</td>
<td>89.6*** [11.2]</td>
</tr>
<tr>
<td>Const</td>
<td>8.5*** [1.6]</td>
<td></td>
</tr>
<tr>
<td>Month FE?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>0.27</td>
<td>0.61</td>
</tr>
<tr>
<td>Adj. R2 (within)</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>1039</td>
<td>1039</td>
</tr>
<tr>
<td>Dates</td>
<td>199601.201012</td>
<td>199601.201012</td>
</tr>
</tbody>
</table>
Table 3: Regressions of MBS returns on OAS and mortgage rates. Panels (a) and (b) show coefficient estimates from OLS regressions of 12-month ($t \rightarrow t + 12$, meaning from the end of month $t$ to the end of month $t + 12$ and reported as “1-year”) MBS returns and annualized 1-month ($t \rightarrow t + 1$, reported as “1-month”) MBS returns from the Bloomberg Barclays index on the Barclays OAS to Treasuries as of the end of month $t$. In panel (a), returns are hedged, while in panel (b), returns are unhedged and net of the risk free rate. Panel (c) shows coefficient estimates from regressions of hedged returns in month $t + 1$ on changes in the Freddie Mac 30-year fixed mortgage rate, denoted “FRM”, from the end of month $t$ to $t + 1$. “ITM” is an indicator for an MBS being in-the-money (i.e. moneyness > 0) as of the end of month $t$. “Market ITM (OTM)” indicates that the balance-weighted average moneyness of all outstanding Fannie Mae MBS is > 0 (< 0). Moneyness is calculated as coupon plus 50 basis points minus the 30-year fixed-rate mortgage rate (from Freddie Mac). Newey-West standard errors (18 lags) in brackets for 1-year returns; robust standard errors (clustered at the month level) for 1-month returns. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

(a) Hedged returns and OAS

<table>
<thead>
<tr>
<th></th>
<th>(1) 1-year</th>
<th>(2) 1-year</th>
<th>(3) 1-month</th>
<th>(4) 1-month</th>
<th>(5) 1-month</th>
<th>(6) 1-month</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAS</td>
<td>1.36***</td>
<td>1.41***</td>
<td>1.75***</td>
<td>1.95***</td>
<td>0.50*</td>
<td>1.15***</td>
</tr>
<tr>
<td></td>
<td>[0.21]</td>
<td>[0.13]</td>
<td>[0.63]</td>
<td>[0.61]</td>
<td>[0.28]</td>
<td>[0.27]</td>
</tr>
<tr>
<td>ΔOAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.37***</td>
<td>-2.33***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.30]</td>
<td>[0.25]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Month FEs? Yes No No Yes No Yes
Adj. R2  0.19  0.70  0.02  0.53  0.56  0.78
Obs.  1990  1990  2076  2076  2064  2064

(b) Unhedged returns and OAS

<table>
<thead>
<tr>
<th></th>
<th>(1) 1-year</th>
<th>(2) 1-year</th>
<th>(3) 1-month</th>
<th>(4) 1-month</th>
<th>(5) 1-month</th>
<th>(6) 1-month</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAS</td>
<td>1.48***</td>
<td>0.80***</td>
<td>1.90**</td>
<td>1.58*</td>
<td>1.04</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>[0.35]</td>
<td>[0.22]</td>
<td>[0.76]</td>
<td>[0.87]</td>
<td>[0.74]</td>
<td>[0.44]</td>
</tr>
<tr>
<td>ΔOAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.85***</td>
<td>-3.32***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.31]</td>
<td>[0.38]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Month FEs? Yes No No Yes No Yes
Adj. R2  0.02  0.78  0.01  0.60  0.11  0.76
Obs.  1990  1990  2076  2076  2064  2064

(c) 1-month hedged returns and mortgage rates

<table>
<thead>
<tr>
<th></th>
<th>(1) Full</th>
<th>(2) Full</th>
<th>(3) Market ITM</th>
<th>(4) Market OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔFRM</td>
<td>-0.58***</td>
<td>-0.74***</td>
<td>-0.39*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.18]</td>
<td>[0.26]</td>
<td>[0.21]</td>
<td></td>
</tr>
<tr>
<td>ΔFRM × ITM</td>
<td>0.92***</td>
<td>1.14***</td>
<td>0.91***</td>
<td>1.57***</td>
</tr>
<tr>
<td></td>
<td>[0.23]</td>
<td>[0.25]</td>
<td>[0.23]</td>
<td>[0.50]</td>
</tr>
</tbody>
</table>

Month FEs? No Yes No No
Adj. R2  0.03  0.55  0.02  0.16
Obs.  2076  2076  1622  454
Table 4: Cross sectional variation of OAS$^Q$ and prepayment risk premia (OAS − OAS$^Q$) on pass-throughs from IO/PO strips. The table shows coefficient estimates from OLS regressions of the OAS$^Q$ (panel a) and the prepayment-risk component in the OAS (panel b) on dummies for different moneyness levels either including or excluding time fixed effects. For each trust, the OAS is computed from the value of the pass-through obtained as the sum of the value of the IO and PO. Refer to Figure 5 for a summary of the relation between OAS and OAS$^Q$ and to Section 4 for more details on the pricing model. Moneyness of a coupon $j$ at time $t$ is defined as $\text{Moneyness}_{j,t} = \text{Coupon}_j + 0.5 - \text{FRMrate}_t$. Moneyness bin $[-.25,.25)$ is the omitted category. Robust standard errors (clustered at the month level) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

(a) Cross section of OAS$^Q$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt; -1.25$</td>
<td>1.5</td>
<td>[3.8]</td>
<td>-1.3</td>
<td>[1.5]</td>
</tr>
<tr>
<td>$[-1.25,-.75)$</td>
<td>-0.2</td>
<td>[2.1]</td>
<td>-2.7**</td>
<td>[1.2]</td>
</tr>
<tr>
<td>$[-.75,-.25)$</td>
<td>-1.4</td>
<td>[1.3]</td>
<td>-2.3**</td>
<td>[0.9]</td>
</tr>
<tr>
<td>$.25, .75)$</td>
<td>-2.3***</td>
<td>[0.7]</td>
<td>-2.5***</td>
<td>[0.8]</td>
</tr>
<tr>
<td>$.75, 1.25)$</td>
<td>-5.2***</td>
<td>[1.6]</td>
<td>-4.0***</td>
<td>[1.5]</td>
</tr>
<tr>
<td>$[1.25, 1.75)</td>
<td>-3.3</td>
<td>[2.5]</td>
<td>-1.4</td>
<td>[2.4]</td>
</tr>
<tr>
<td>Const</td>
<td>13.8***</td>
<td>[1.6]</td>
<td>13.1***</td>
<td>[0.9]</td>
</tr>
</tbody>
</table>

(b) Cross section of OAS - OAS$^Q$

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt; -1.25$</td>
<td>5.4***</td>
<td>[1.9]</td>
<td>6.8***</td>
<td>[2.2]</td>
</tr>
<tr>
<td>$[-1.25,-.75)$</td>
<td>2.9***</td>
<td>[0.9]</td>
<td>5.2***</td>
<td>[1.2]</td>
</tr>
<tr>
<td>$[-.75,-.25)$</td>
<td>-1.0***</td>
<td>[0.4]</td>
<td>0.7</td>
<td>[0.7]</td>
</tr>
<tr>
<td>$.25, .75)$</td>
<td>4.9***</td>
<td>[0.4]</td>
<td>4.1***</td>
<td>[0.5]</td>
</tr>
<tr>
<td>$.75, 1.25)$</td>
<td>13.3***</td>
<td>[1.0]</td>
<td>12.4***</td>
<td>[1.1]</td>
</tr>
<tr>
<td>$\geq 2.25$</td>
<td>34.2***</td>
<td>[5.0]</td>
<td>27.5***</td>
<td>[4.3]</td>
</tr>
<tr>
<td>Const</td>
<td>-1.4***</td>
<td>[0.2]</td>
<td>-1.0</td>
<td>[0.7]</td>
</tr>
</tbody>
</table>
Table 5: Time-series regressions of OAS, OAS\textsuperscript{Q} and prepayment risk premia on pass-throughs from IO/PO strips. The table shows coefficient estimates from OLS regressions of spreads reported in the top row of each panel on the Baa-Aaa corporate spread, MBS issuance to equity of brokers and dealers, average moneyness in the MBS universe squared, spreads on agency debt to swaps and dealer disagreement about prepayment speeds. For each trust, the OAS is computed from the value of the pass-through obtained as the sum of the value of the IO and PO. Refer to Figure 5 for a summary of the relation between OAS and OAS\textsuperscript{Q} and to Section 4 for more details on the pricing model. See Section 5.2 for more detail on variable construction. Data is at monthly frequency and measured as of month-end. Columns (1) to (3) cover the entire sample period 1995-2010, while columns (4) to (6) exclude August 2007 to December 2010 (financial crisis and aftermath). Regressors are standardized to have zero mean and unit standard deviation. Newey-West standard errors (6 lags) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

<table>
<thead>
<tr>
<th></th>
<th>(1) OAS</th>
<th>(2) OAS\textsuperscript{Q}</th>
<th>(3) OAS - OAS\textsuperscript{Q}</th>
<th>(4) OAS</th>
<th>(5) OAS\textsuperscript{Q}</th>
<th>(6) OAS - OAS\textsuperscript{Q}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baa - Aaa</td>
<td>15.5***</td>
<td>13.0***</td>
<td>2.5***</td>
<td>12.0***</td>
<td>11.7**</td>
<td>0.2</td>
</tr>
<tr>
<td>MBS Issuance / BD Equity</td>
<td>5.8***</td>
<td>4.3***</td>
<td>1.5*</td>
<td>4.6***</td>
<td>2.3</td>
<td>2.2</td>
</tr>
<tr>
<td>Mean Moneyness\textsuperscript{2}</td>
<td>2.5</td>
<td>-3.8**</td>
<td>6.3***</td>
<td>4.1**</td>
<td>-1.1</td>
<td>5.1***</td>
</tr>
<tr>
<td>Agency Debt Spreads</td>
<td>5.9***</td>
<td>7.3***</td>
<td>-1.4</td>
<td>6.2***</td>
<td>8.0***</td>
<td>-1.7</td>
</tr>
<tr>
<td>Disagreement</td>
<td>-1.7</td>
<td>-3.1</td>
<td>1.4*</td>
<td>3.0***</td>
<td>0.8</td>
<td>2.2***</td>
</tr>
<tr>
<td>Const</td>
<td>17.2***</td>
<td>11.8***</td>
<td>5.4***</td>
<td>15.7***</td>
<td>11.4***</td>
<td>4.3***</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>0.69</td>
<td>0.62</td>
<td>0.59</td>
<td>0.67</td>
<td>0.47</td>
<td>0.50</td>
</tr>
<tr>
<td>Obs.</td>
<td>192</td>
<td>192</td>
<td>192</td>
<td>151</td>
<td>151</td>
<td>151</td>
</tr>
</tbody>
</table>
Appendix

A  MBS cash flows

This Appendix provides detail on the calculation of cash flows for an MBS, which are used to compute mortgage spreads as described in Section 2. Consider a fixed-rate MBS with an original balance of $1, and let $\theta_t$ be the “factor,” or remaining balance relative to origination, at date $t$. In level-payment fixed-rate mortgages, the principal is repaid gradually rather than with a bullet payment at maturity and the borrower makes fixed payments every month. Denote the loan maturity measured in months by $T$ (at the pool level, this is referred to as weighted average maturity, or WAM). Let $k$ be the monthly installment from the borrower to the servicer, $w$ the interest rate on the loan (or weighted average coupon, WAC, at the pool level), and $c$ the coupon paid to investors. The difference between the loan and coupon rates is earned by servicers or by the guaranteeing agency. To compute the fixed payment $k$ note that, net of this payment, the loan balance absent any prepayment, denoted $\tilde{\theta}_t$ grows at rate $(1 + w)$, or:

$$\tilde{\theta}_t = \left(\frac{1 + w}{1 + w} \right)^T - 1 - k. \quad (A.1)$$

Solving for $\tilde{\theta}_T = 0$, it then follows that $k = \left(\frac{w(1 + w)^T}{(1 + w)^T - 1}\right)$. The evolution of the loan balance $\theta_t$ allowing for early prepayment generalizes equation (A.1). After accounting for loan amortization and unscheduled principal payments, the factor evolves according to:

$$\theta_t = (1 - SMM_t - 1) - k \hat{\theta}_t, \quad (A.2)$$

where $w$ is the interest rate on the loan (or weighted average coupon, WAC) and $k$ is the constant monthly payment composed of the scheduled principal and interest payments. SMM$_t$ is the “single month mortality,” or the fraction of the remaining balance that was prepaid in month $t$ due to unscheduled principal payments, and $\hat{\theta}_t$ is the cumulated fraction of unit principal that has not prepaid since the inception of the mortgage, which is also known as the survival factor, $\hat{\theta}_t = \prod_{s=0}^{t-1} (1 - SMM_s)$. It then follows that $\theta_t = \hat{\theta}_t \times \tilde{\theta}_t$.\footnote{The prepayment speed is often reported in annualized terms, known as the “conditional prepayment rate” or CPR$_t = 1 - (1 - SMM_t)^{12}$.} Given prepayment rates, cash flows passed through to investors per unit of principal are:

$$X_t = (\theta_{t-1} - \theta_t) + c \theta_{t-1}, \quad (A.3)$$

where the principal payment is equal to the decline in principal $(\theta_{t-1} - \theta_t)$ and the coupon payment from the borrower to the investor net of the servicing and agency guarantee fees is $c \theta_{t-1}$.1
B Data

B.1 TBA sample and data-quality filters

The sample used for the analysis in Sections 2.2 and 2.3 spans 1996 to 2010, reflecting limited availability of data on MBS in the TBA market prior to 1996, and a limited liquidity in IO/PO strips, which we use later in the paper to decompose the OAS, post 2010. The OAS we use are the average (within each coupon/month) across six dealers from which we collect data. We do not necessarily have spreads for all dealers on the same coupons on each day. In addition, some of the dealers enter our data only after 1996. We clean each dealer’s data to prevent spreads from being influenced by stale prices. To do so, we check whether a price for a coupon is unchanged relative to the previous day. If it is, and if the 10-year Treasury yield changed by 3 basis points or more on the same day (so we expect MBS prices to change), we drop the price and the corresponding spread. If the price is not constant, but had been constant more than twice in the same calendar month on days when the Treasury yield moves, we similarly drop it.

Descriptive statistics of the monthly coupon-level sample used for the analysis in Sections 2.2 and 2.3 are provided in Table A-1. The coupons in the sample change over time (as illustrated in Figure 1 in the main text), so that the data form an unbalanced panel.

Table A-1: Descriptive statistics of TBA sample. Based on Fannie Mae pass-through coupons from 1996-2010 for which we have OAS quotes from at least one dealer, subject to the data quality filters described in Section B.1. Data frequency is monthly; spreads and moneyness (= coupon plus 0.5 minus current FRM rate) are as of month-end. RPB stands for the total remaining principal balance in a coupon. Factor means current face value divided by issuance amount. Data on RPBs, factors, and weighted average loan ages are from eMBS.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>p5</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p95</th>
<th>Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAS (basis points, annual)</td>
<td>23.12</td>
<td>50.66</td>
<td>-16.45</td>
<td>-2.27</td>
<td>9.05</td>
<td>28.03</td>
<td>120.00</td>
<td>1532</td>
</tr>
<tr>
<td>Coupon (percent)</td>
<td>6.53</td>
<td>1.49</td>
<td>4.00</td>
<td>5.50</td>
<td>6.50</td>
<td>7.50</td>
<td>9.00</td>
<td>1532</td>
</tr>
<tr>
<td>Moneyness (percent)</td>
<td>0.56</td>
<td>1.40</td>
<td>-1.64</td>
<td>-0.58</td>
<td>0.51</td>
<td>1.60</td>
<td>2.93</td>
<td>1532</td>
</tr>
<tr>
<td>RPB (2009 USD, million)</td>
<td>126928.83</td>
<td>151009.18</td>
<td>590.43</td>
<td>9091.44</td>
<td>66862.53</td>
<td>194624.09</td>
<td>443893.38</td>
<td>1532</td>
</tr>
<tr>
<td>Factor of most recent vintage</td>
<td>0.89</td>
<td>0.21</td>
<td>0.26</td>
<td>0.91</td>
<td>0.98</td>
<td>0.99</td>
<td>1.00</td>
<td>1530</td>
</tr>
<tr>
<td>Weighted average loan age (months)</td>
<td>47.44</td>
<td>33.51</td>
<td>8.67</td>
<td>25.37</td>
<td>38.36</td>
<td>58.93</td>
<td>128.27</td>
<td>1532</td>
</tr>
</tbody>
</table>

B.2 Stripped MBS data-quality filters

We start with daily price quotes from a large dealer for the period 1995 to 2010, and then clean these data using the following steps:

1. Remove/correct obvious outliers (such as prices of 0 or a few instances where IO and PO prices were inverted).
2. Remove prices that are stale (defined as a price that does not change from previous day despite a change in the 10-year yield of more than 3 basis points). In case of smaller yield
changes, we check the previous 10 days and remove a price if there were more than two instances of stale prices on that security over that period.

3. For a subsample of trusts and months (starting in June 1999), we also have price quotes from two additional dealers. When available, we compare these prices (their average if both are available) to the price quoted by our dealer. When they are more than 5% apart, or if the overall range of price quotes is larger than 0.1 times the average price, we do not use our price quote in the analysis. This applies to about 10% of our price quotes.

4. Only retain trusts for which we have both the IO and PO strips, and which we can link to data on the underlying pool of mortgages (from eMBS). This restriction eliminates IO strips backed by excess servicing rights, for instance.

5. Only retain trusts for which the price on the recombined pass-through \(= P_{PO} + P_{IO} \) is within $2 of the TBA price of the corresponding coupon. (We also drop trusts if on that day we do not have a clean TBA price for the corresponding coupon.)

6. Only retain trusts with a factor \(=\) current face value divided by issuance amount) of more than 5%.

7. Only retain trusts that we can match to a Bloomberg prepayment survey with the same coupon and absolute differences in WAC and WAM smaller than 0.3 percentage points and 60 months, respectively. (This affects almost exclusively observations before 2003, as we do not have individual vintages in the survey in the early years.)

Following these steps, the sample includes 3713 trust-month observations, or about 19 per month. The year with the lowest number of observations is 1999, where we have an average of 10 trusts per month, while after 2005 we have at least 20 trusts in all but one month.

C OAS smile: Robustness

This Appendix provides additional evidence on the smile-shaped pattern of the OAS in the cross section of TBA coupons, which is a key empirical result as discussed in Section 2.3. We show robustness to dropping periods of financial stress, stability across sub-periods, and regression estimates when excluding outlier observations, MBS coupons with relatively low balances, or coupons that may be “burned out” (as indicated by a low remaining factor).

One potential concern about the OAS pattern shown in panel (b) of Figure 2 is that it may be unduly influenced by periods of financial stress. Panel (a) of Figure A-1 replicates Figure 2 when excluding the post-July 2007 period, and the period from September 1998 to January 1999, when liquidity in many U.S. fixed income markets dried up following the failure of Long-Term Capital Management (LTCM). The figure shows that the distinctive smile pattern remains present even without these periods. Panel (b) of Figure A-1 provides evidence on sub-sample stability by
providing “binned” scatter plots for three equally sized periods covering five years each. While the smile pattern is most distinctive during the last five years of the sample, it is also clearly visible in the earlier time periods.

We next present additional robustness checks of the regression evidence shown in Table 2 of the main text, focusing on the specifications that explore within-month variation in OAS across coupons. Table 2 showed that the smile pattern is robust to restricting the sample to certain time periods, either based on dates or based on the moneyness of the market as a whole. We now instead drop observations based on the characteristics of the underlying coupons, or if their OAS is an outlier. Column (1) of Table A-2 displays, for reference, the results of our baseline specification with month fixed effects (column 2 in Table 2(a)). In column (2), we drop from the sample all observations where the absolute level of moneyness exceeds 2, or where the OAS exceeds 100 basis points. The resulting coefficients confirm that the smile pattern is not driven by extreme observations. The final two columns instead restrict the sample based on characteristics of the pools underlying each coupon, to ensure that the smile pattern is not due just to small pools. In column (3), we restrict the included coupons to be those from the original sample with above-median remaining principal balance (in 2009 dollars); this cutoff is $67 billion. We see that the coefficients on the most extreme bins are slightly reduced relative to column (1), but otherwise the smile pattern remains very similar. Finally, in column (4) we restrict the sample to include only those coupons where the most recently issued origination vintage (by year) has a remaining factor of at least 0.8 (meaning at most 20% of the original pool has prepaid). This excludes the coupons with the most "burned out" pools. This restriction has little effect on the estimated coefficients.

Table A-2: Cross section of OAS on TBA coupons: Robustness. Coefficient estimates from OLS regression of the OAS on different moneyness level bins, with calendar month fixed effects. Robust standard errors (clustered at the month level) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

<table>
<thead>
<tr>
<th></th>
<th>(1) Full</th>
<th>(2)</th>
<th>(3) RPB&gt; med.</th>
<th>(4) Factor&gt;0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt; -1.75$</td>
<td>18.0***[2.3]</td>
<td>16.5***[2.6]</td>
<td>5.9***[2.3]</td>
<td>17.9***[2.2]</td>
</tr>
<tr>
<td>$[-1.75, -1.25)$</td>
<td>10.4***[1.6]</td>
<td>9.8***[1.3]</td>
<td>3.7***[1.1]</td>
<td>9.5***[1.5]</td>
</tr>
<tr>
<td>$[-1.25, -0.75)$</td>
<td>5.3***[1.1]</td>
<td>4.9***[0.8]</td>
<td>-0.4[0.9]</td>
<td>5.8***[1.0]</td>
</tr>
<tr>
<td>$[-0.75, -0.25)$</td>
<td>-0.1[0.4]</td>
<td>0.1[0.4]</td>
<td>-1.1[0.7]</td>
<td>-0.1[0.4]</td>
</tr>
<tr>
<td>$[0.25, 0.75)$</td>
<td>1.5***[0.5]</td>
<td>1.2**[0.5]</td>
<td>1.2**[0.7]</td>
<td>1.5***[0.5]</td>
</tr>
<tr>
<td>$[0.75, 1.25)$</td>
<td>6.4***[1.2]</td>
<td>5.1***[1.0]</td>
<td>4.7***[1.4]</td>
<td>6.3***[1.2]</td>
</tr>
<tr>
<td>$[1.25, 1.75)$</td>
<td>15.0***[2.1]</td>
<td>11.0***[1.7]</td>
<td>12.8***[2.8]</td>
<td>15.7***[2.2]</td>
</tr>
<tr>
<td>$[2.25, 2.75)$</td>
<td>33.9***[5.5]</td>
<td>38.9***[7.5]</td>
<td>48.6***[9.4]</td>
<td></td>
</tr>
<tr>
<td>$\geq 2.75$</td>
<td>85.4***[10.9]</td>
<td>51.8***[2.6]</td>
<td>91.6***[15.2]</td>
<td></td>
</tr>
</tbody>
</table>

| Month FEs? | Yes | Yes | Yes | Yes |
| Adj. R2 | 0.61 | 0.68 | 0.88 | 0.76 |
| Adj. R2 (within) | 0.31 | 0.15 | 0.31 | 0.32 |
| Obs. | 1532 | 1215 | 766 | 1314 |
Figure A-1: OAS smile: robustness and subsamples. Panel (a) displays a scatterplot and a local smoother of the cross-sectional variation in OAS for MBS coupons with remaining principal balance (in 2009 dollars) of $100 million or more, excluding from the sample the months September 1998 to January 1999 (LTCM turmoil) and August 2007 to December 2010 (financial crisis and aftermath). Panel (b) displays “binned” scatterplots for three five-year subsamples, where each dot represents mean OAS for one of ten equal-sized moneyness bins. The lines represent a local smoother fitted to the underlying data (i.e., each coupon’s OAS, not just the moneyness bin averages).

(a) Excluding periods of financial stress

(b) Five-year subsamples
D Expected return derivations

In this Appendix, we derive the relation between OAS and future excess returns (2.3) and the relation between contemporaneous excess returns and changes in OAS shown in (2.4). We finally show that the expected hedged return under the interest-rate-risk-neutral measure \( Q \) equals the expected hedged return under the physical measure \( P \).

**Derivation of (2.3).** We begin by showing that the OAS can be approximated as a weighted average of future excess returns. Approximating the right hand side of (2.2) around 0 excess return, we obtain

\[
\frac{\partial}{\partial \mathbb{E}^Q_{t+j-1}[r_{x,t+j}]} \prod_{j=1}^{k} \left(1 + r_{t+j} + \mathbb{E}^Q_{t+j-1}[r_{x,t+j}]\right)^{-1} \bigg|_{\mathbb{E}^Q_{t+j-1}[r_{x,t+j}]=0} = -\frac{1}{1 + r_{t+j}} \prod_{j=1}^{k} \frac{1}{1 + r_{t+j}},
\]

so that (2.2) becomes

\[
P_t \approx \mathbb{E}^Q_t \left[ \sum_{k=1}^{T-t} X_k \frac{1}{\prod_{j=1}^{k} \left(1 + r_{t+j}\right)} \left(1 - \sum_{l=1}^{k} \mathbb{E}^Q_{t+l-1}[r_{x,t+l}] \frac{1}{1 + r_{t+l}}\right) \right].
\]

Similarly, approximating the right hand side of (2.1) around 0 OAS, we can express

\[
\frac{\partial}{\partial \text{OAS}_t} \prod_{j=1}^{k} \left(1 + r_{t+j} + \text{OAS}_t\right)^{-1} \bigg|_{\text{OAS}_t=0} = -\frac{1}{1 + r_{t+l}} \prod_{j=1}^{k} \frac{1}{1 + r_{t+j}},
\]

so that (2.1) becomes

\[
P_t \approx \mathbb{E}^Q_t \left[ \sum_{k=1}^{T-t} X_k \frac{1}{\prod_{j=1}^{k} \left(1 + r_{t+j}\right)} \left(1 - \sum_{l=1}^{k} \text{OAS}_t \frac{1}{1 + r_{t+l}}\right) \right].
\]

Comparing these two expressions, we obtain

\[
\text{OAS} \mathbb{E}^Q_t \left[ \sum_{k=1}^{T-t} X_k \frac{1}{\prod_{j=1}^{k} \left(1 + r_{t+j}\right)} \frac{1}{1 + r_{t+l}} \right] = \mathbb{E}^Q_t \left[ \sum_{k=1}^{T-t} X_k \frac{1}{\prod_{j=1}^{k} \left(1 + r_{t+j}\right)} \sum_{l=1}^{k} \mathbb{E}^Q_{t+l-1}[r_{x,t+l}] \frac{1}{1 + r_{t+l}} \right],
\]

and rearranging we obtain (2.3), where weights are:

\[
w_{k,t} = \sum_{l=k}^{T-t} \frac{X_{t+l}}{\prod_{k=1}^{l} \left(1 + r_{t+j}\right)} \frac{1}{1 + r_{t+k}} \left[ \sum_{k=1}^{T-t} \frac{1}{1 + r_{t+k}} \sum_{l=k}^{T-t} \frac{X_{t+l}}{\prod_{j=1}^{l} \left(1 + r_{t+j}\right)} \right].
\]
Derivation of (2.4). To understand the link between realized excess returns and changes in the OAS, from (2.1), the OAS at date \( t + 1 \) is
\[
P_{t+1} = \mathbb{E}_{t+1}^{Q_t} \left[ \frac{\sum_{k=1}^{T-t-1} X_{t+1+k}}{\prod_{j=1}^{k} (1 + r_{t+1+j} + OAS_t)} \right].
\]

Applying a first order Taylor approximation around the OAS remaining the same between dates \( t \) and \( t + 1 \), the price next period can be approximated by
\[
P_{t+1} \approx \mathbb{E}_{t+1}^{Q_t} \left[ \frac{\sum_{k=1}^{T-t-1} X_{t+1+k}}{\prod_{j=1}^{k} (1 + r_{t+1+j} + OAS_t)} \right] + \frac{\partial P_{t+1}}{\partial OAS} \Delta OAS_{t+1},
\]
where \( \Delta OAS_{t+1} = OAS_{t+1} - OAS_t \) is the change in the OAS. Thus, the one period return to holding the MBS can be expressed as
\[
1 + r_{t+1} + rx_{t+1} = \frac{P_{t+1} + X_{t+1}}{P_t} \approx \frac{\mathbb{E}_{t+1}^{Q_t} \left[ \frac{\sum_{k=1}^{T-t-1} X_{t+1+k}}{\prod_{j=1}^{k} (1 + r_{t+1+j} + OAS_t)} \right] + \frac{\partial P_{t+1}}{\partial OAS} \Delta OAS_{t+1}}{P_t} + \frac{1}{P_t} \frac{\partial P_{t+1}}{\partial OAS} \Delta OAS_{t+1},
\]
where the last equality uses the fact that \( 1 + r_{t+1} + OAS_t \) and \( X_{t+1} \) are known at date \( t + 1 \), so that we can bring \( X_{t+1} \) inside the date \( t + 1 \) expectations operator and factor \( 1 + r_{t+1} + OAS_t \) from the date \( t + 1 \) expectation. Since the ratio of expectations is approximately 1, we obtain (2.4), where duration \( D_t \) is defined as \( -\frac{1}{P_t} \frac{\partial P_{t+1}}{\partial OAS} \). As discussed in the main text, this expression is the analog to what Shiller et al. (1983) derive in the context of coupon-bearing Treasuries. This approximate relation between returns and changes in OAS is also employed by practitioners (see Lehman Brothers, 2008, for example).

Expected excess returns under \( P \) and \( Q_t \). Using the fund separation theorem (see e.g. Merton, 1972), we can represent the return on the MBS as
\[
rx_{t+1} = \sum_{j=1}^{N_p} \beta^i_{r,jt} z^j_{t+1} + \sum_{j=1}^{N_p} \beta^j_{r,lt} p^j_{t+1} + \epsilon_{t+1}, \tag{D.1}
\]
\[
= \sum_{j=1}^{N_p} \beta^j_{r,lt} p^j_{t+1} + \epsilon_{t+1},
\]
where \( \{z_{i,t+1}^j\}_{j=1}^{N_r} \) are the one-period excess returns on a set of portfolios that spans interest rate uncertainty, and \( \{p_{i,t+1}^j\}_{j=1}^{N_p} \) are the one-period excess returns on a set of portfolios that spans the prepayment risk that is orthogonal to interest rates. Here, \( V_t \) and \( C_t \) represent conditional variance and covariance, respectively. Thus, (D.1) represents excess returns on the MBS in terms of its beta with respect to interest rate uncertainty spanning portfolios, non-interest rate prepayment uncertainty spanning portfolios and idiosyncratic risk, \( \epsilon_{t+1} \).

The return on the hedged portfolio is then given by

\[
rx_{t+1} - \sum_{j=1}^{N_r} \beta_{i,t}^j z_{t+1}^j = \sum_{j=1}^{N_p} \beta_{i,t}^j p_{t+1}^j + \epsilon_{t+1}.
\]

Then the expected hedged return under the interest-rate-neutral measure \( Q_r \) is given by

\[
E_t^{Q_r} \left[ rx_{t+1} - \sum_{j=1}^{N_r} \beta_{i,t}^j z_{t+1}^j \right] = E_t^{Q_r} \left[ \sum_{j=1}^{N_p} \beta_{i,t}^j p_{t+1}^j + \epsilon_{t+1} \right] = E_t \left[ M_{r,t+1} \left( \sum_{j=1}^{N_p} \beta_{i,t}^j p_{t+1}^j + \epsilon_{t+1} \right) \right],
\]

where \( M_{r,t+1} \) is the change of measure between \( P \) and \( Q_r \). Since \( \{z_{i,t+1}^j\}_{j=1}^{N_r} \) span the interest rate uncertainty, \( M_{r,t+1} \) can be represented as a portfolio of \( \{z_{t+1}^j\}_{j=1}^{N_r} \)

\[
M_{r,t+1} = m_{0t} + \sum_{j=1}^{N_r} m_j z_{t+1}^j,
\]

with 1 = \( m_{0t} + \sum_{j=1}^{N_r} m_j E_t \left[ z_{t+1}^j \right] \). Thus

\[
E_t^{Q_r} \left[ rx_{t+1} - \sum_{j=1}^{N_r} \beta_{i,t}^j z_{t+1}^j \right] = E_t \left[ \left( m_{0t} + \sum_{j=1}^{N_r} m_j z_{t+1}^j \right) \left( \sum_{j=1}^{N_p} \beta_{i,t}^j p_{t+1}^j + \epsilon_{t+1} \right) \right]
\]

\[
= E_t \left[ m_{0t} + \sum_{j=1}^{N_r} m_j z_{t+1}^j \right] E_t \left[ \sum_{j=1}^{N_p} \beta_{i,t}^j p_{t+1}^j + \epsilon_{t+1} \right]
\]

\[
= E_t \left[ \sum_{j=1}^{N_p} \beta_{i,t}^j p_{t+1}^j + \epsilon_{t+1} \right],
\]

where the second equality uses the fact that \( z_{t+1}^j \) and \( p_{t+1}^j \) are uncorrelated and the third equality uses the fact that the change of measure \( M_{r,t+1} \) is a martingale.
E Additional details on model

In this Appendix, we describe more formally, and under more general assumptions on the evolution of the term structure of interest rates, the simple framework presented in Section 3 in the main text. We show that the simple intuitions for the sources of risk premia reflected in the OAS carry through even under these more general assumptions.

We consider a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \(\Omega\) is the set of states of nature with generic element \(\omega\), \(\mathcal{F}\) is the \(\sigma\)-algebra of observable events and \(\mathbb{P}\) is a probability measure on \((\Omega, \mathcal{F})\). We assume that the uncertainty in the economy is generated by a \(N\)-dimensional Brownian motion \(Z_t\) and a Poisson process \(J_t\), with \(J_t\) denoting the number of jumps up to time \(t\). We further partition the Brownian motion \(Z_t\) into \(N_r\) shocks that affect the risk-free interest rate in the economy and \(N - N_r\) shocks that do not: \(Z_t = \begin{bmatrix} Z_{t}^{r} & Z_{t}^{\gamma} \end{bmatrix}'\). The Poisson process \(J_t\) captures liquidity shocks to the MBS market and arrives with intensity \(\mu_t\); if a jump occurs at date \(t\), the securities written on pool \(j\) lose a fraction \(\alpha_{jt}\) of their market value.

Consider a mortgage pool \(j\) with coupon rate \(c_j\), remaining principal balance \(\theta_{jt}\) at time \(t\) and maturity date \(T_j\). As in Appendix A, let \(\bar{\theta}_{jt}\) be the remaining loan balance absent unscheduled prepayment at date \(t\) and \(\bar{c}_j\) is the fixed interest rate paid by the mortgagor (usually \(\bar{c}_j = c_j + 50\text{bps}\)). Analogously to (A.1), in continuous time, this loan balance evolves as

\[
\frac{\partial \bar{\theta}_{jt}}{\partial t} = \bar{c}_j \bar{\theta}_{jt} - k,
\]

where \(k\) is, as before, the fixed payment from the borrower to the servicer. Solving for \(\bar{\theta}_{jT} = 0\) and \(\bar{\theta}_{jt} = \bar{\theta}_{j0}\), we obtain that the loan balance evolves as

\[
\bar{\theta}_{jt} = \bar{\theta}_{j0} e^{\bar{c}_j T_j} - e^{\bar{c}_j t} e^{\bar{c}_j T_j - 1},
\]

where \(\bar{\theta}_{j0}\) is the balance at origination and the fixed payment \(k\) is \(\left(\frac{\bar{c}_j}{\bar{c}_j e^{\bar{c}_j T_j} - 1}\right)\). Denote by \(S^*_jt\) the cumulative unscheduled prepayments, so that the remaining principal balance at date \(t\) is

\[
\theta_{jt} = e^{-S^*_jt} \bar{\theta}_{jt}.
\]

We assume that the cumulative unscheduled prepayments can be represented as

\[
S^*_jt = \int_0^t s^*_jt u du,
\]

where \(s^*_jt\) is a function of the interest rate incentive \(\bar{c}_j - \rho_t\), where \(\rho_t\) is the interest rate on new loans, and a vector of parameters \(\gamma_t = (\gamma_{1t}, \gamma_{2t}, \ldots, \gamma_{T_t})\):

\[
s^*_jt = f (\gamma_t, \bar{c}_j - \rho_t).
\]
Finally, let \( S_{jt} \) be the full (that is, the sum of both scheduled and unscheduled) cumulative prepayment, so that

\[
\theta_{jt} = e^{-S_{jt}} \theta_{j0},
\]

where \( \theta_{j0} \equiv \tilde{\theta}_{j0} \). With this notation, the cash-flows received by investors in an IO strip on pool \( j \) are \( dX_{jt,IO} = c_j \theta_{jt} dt \), by investors in a PO strip on pool \( j \) are \( dX_{jt,PO} = -d\theta_{jt} \), and by investors in the pass-through on pool \( j \) (the combination, denoted “PT,” of IO and PO) are \( dX_{jt,PT} = c_j \theta_{jt} dt - d\theta_{jt} \).

Under the assumption of no arbitrage, there exists a stochastic discount factor that prices all the assets in the economy, given by

\[
\frac{dM_t}{M_t} = \exp \left( -r_t dt - \pi_{rt}' dZ_{rt} - \pi_{r\gamma}' dZ_{r\gamma} - \pi_{lt} dJ_t \right),
\]

where \( r_t \) is the equilibrium risk-free rate, \( \pi_{rt} \) is the vector of the prices of risk associated with innovations affecting the term structure of interest rates in the economy (including \( \rho_t \)), \( \pi_{r\gamma} \) is the vector of the prices of risk associated with the innovations to the non-interest rate factors, and \( \pi_{lt} \) is the price of risk associated with the liquidity shocks \( dJ_t \). Associated with the stochastic discount factor is the risk-neutral measure \( Q \). In addition to the risk-neutral measure, we consider two intermediate changes of measure: one that is risk-neutral with respect to the interest rate shocks only, \( Q_r \), and one that is risk-neutral with respect to all Brownian shocks, \( Q_{r\gamma} \). These two measures are defined, respectively, by their Radon-Nikodym derivatives with respect to the physical measure \( \mathbb{P} \):

\[
\frac{dQ_r}{d\mathbb{P}} = \exp \left( -\frac{1}{2} |\pi_{rt}|^2 dt - \pi_{rt}' dZ_{rt} \right)
\]

\[
\frac{dQ_{r\gamma}}{d\mathbb{P}} = \exp \left( -\frac{1}{2} \left( |\pi_{rt}|^2 + |\pi_{r\gamma}|^2 \right) dt - \pi_{rt}' dZ_{rt} - \pi_{r\gamma}' dZ_{r\gamma} \right).
\]

We can now introduce our two measures of the option-adjusted spread. Let \( P_{k,jt} \) be the market price of a security of type \( k \in \{IO, PO, PT\} \) on pool \( j \) at time \( t \). The market price is given by

\[
P_{k,jt} = \mathbb{E}_t \left[ \int_0^{+\infty} \frac{M_{t+s}}{M_t} dX_{k,jt+s} \right] = \mathbb{E}_t^Q \left[ \int_0^{+\infty} \exp \left( -\int_0^s r_{t+u} du \right) dX_{k,jt+s} \right],
\]

where the second equality uses the definition of the risk-neutral measure \( Q \). The traditional OAS is defined implicitly through

\[
P_{k,jt} = \mathbb{E}_t^{Q_t} \left[ \int_0^{+\infty} \exp \left( -\int_0^s r_{t+u} du - OAS_{k,jt}s \right) dX_{k,jt+s} \right], \tag{E.1}
\]
while the prepayment-risk-neutral OAS, or $OAS^Q$, is defined implicitly through

$$P_{k,jt} = E_t^{Q_t} \left[ \int_0^{+\infty} \exp \left( - \int_0^s \left( r_{t+u} + OAS_{k,jt+u} \right) du \right) dX_{k,jt+s} \right]. \quad (E.2)$$

As we show in Appendix D, the OAS is a weighted average of expected excess returns on the security. We follow the same logic here and replace the average OAS over the security’s lifetime with a stream of instantaneous OAS, so that (E.1) and (E.2) become

$$P_{k,jt} = E_t^{Q_t} \left[ \int_0^{+\infty} \exp \left( - \int_0^s \left( r_{t+u} + OAS_{k,jt+u}^{Q_t} \right) du \right) dX_{k,jt+s} \right].$$

Differentiating both sides, we obtain

$$-E_t^{Q_t} \left[ dP_{k,jt} \right] = - \left( r_t + OAS_{k,jt} \right) P_{k,jt} - dt + E_t^{Q_t} \left[ dX_{k,jt} \right]$$

$$-E_t^{Q_t} \left[ dP_{k,jt} \right] = - \left( r_t + OAS_{k,jt}^{Q_t} \right) P_{k,jt} - dt + E_t^{Q_t} \left[ dX_{k,jt} \right]$$

$$-E_t^{Q_t} \left[ dP_{k,jt} \right] = -r_t P_{k,jt} - dt + E_t^{Q_t} \left[ dX_{k,jt} \right].$$

Comparing the three expressions for the equilibrium price, we can represent $OAS_{k,jt}$ and $OAS_{k,jt}^{Q_t}$ as

$$OAS_{k,jt} P_{k,jt} - dt = E_t^{Q_t} \left[ dP_{k,jt} + dX_{k,jt} \right] - E_t^{Q_t} \left[ dP_{k,jt} + dX_{k,jt} \right]$$

$$OAS_{k,jt}^{Q_t} P_{k,jt} - dt = E_t^{Q_t} \left[ dP_{k,jt} + dX_{k,jt} \right] - E_t^{Q_t} \left[ dP_{k,jt} + dX_{k,jt} \right].$$

Applying Ito’s lemma, we obtain

$$dP_{k,jt} = \frac{\partial P_{k,jt}}{\partial t} dt + \frac{\partial P_{k,jt}}{\partial \rho_t} d\rho_t + \frac{\partial P_{k,jt}}{\partial \gamma_t} d\gamma_t + \frac{1}{2} \frac{\partial^2 P_{k,jt}}{\partial \rho_t^2} (d\rho_t)^2 + \frac{1}{2} tr \left( \frac{\partial^2 P_{k,jt}}{\partial \gamma_t^2} \right) \left( d\gamma_t \right)^2$$

$$+ \frac{\partial P_{k,jt}}{\partial \rho_t} \langle d\rho_t, d\gamma_t \rangle + \left[ P_{k,jt} - P_{k,jt-} \right]$$

$$d\theta_{jt} = -\tilde{c}_j e^{\delta_j \theta_{jt}} - e^{-\delta_j \theta_{jt}} \theta_{jt} dt - \theta_{jt} s_j^r dt.$$
take the physical evolution of the jump process, so that the expectation of the jump component of the price are equivalent under the three measures. Thus, the expressions for the \(OAS_{k,jt}^P\) and \(OAS_{k,jt}^Q\) become

\[
OAS_{k,jt}^P P_{k,jt} \cdot dt = \left( E_t - E_t^Q \right) \left[ \partial P_{k,jt} \over \partial \gamma_t \right] d\gamma_t + \left( E_t^Q - E_t^P \right) [dJ_t] \alpha_j P_{k,jt}^-. \tag{E.3}
\]

\[
OAS_{k,jt}^Q P_{k,jt} \cdot dt = \left( E_t^Q - E_t^P \right) [dJ_t] \alpha_j P_{k,jt}^- . \tag{E.4}
\]

Thus, the traditional OAS combines the prepayment risk premium

\[
\frac{1}{P_{k,jt}^-} \left( E_t - E_t^Q \right) \left[ \partial P_{k,jt} \over \partial \gamma_t \right] d\gamma_t ,
\]

with the liquidity risk premium

\[
\frac{1}{dt} \left( E_t^Q - E_t^P \right) [dJ_t] \alpha_j.
\]

The prepayment-risk-neutral OAS, on the other hand, is only composed of the liquidity risk premium. Under the assumption that all securities written on pool \(j\) lose the same fraction \(\alpha_{jt}\) of their value when a liquidity jump occurs, \(OAS^Q\) is equal for all securities written on the pool, since the liquidity risk premium depends only on the difference between the risk-neutral and the physical probabilities of the jump occurring, \(1 \over dt \left( E_t^Q - E_t^P \right) [dJ_t]\), and the loss in valuation due to the jump

\[
OAS_{IO,jt}^Q = OAS_{PO,jt}^Q = OAS_{PT,jt}^Q = \frac{1}{dt} \left( E_t^Q - E_t^P \right) [dJ_t] \alpha_{jt} . \tag{E.5}
\]

Thus, Proposition 4.1 holds.

Consider now the example in Section 3. In Section 3, we assume that the risk-free rate and mortgage rate are constant and equal to each other, so that \(r_t = \rho_t = r\), that, under the physical measure \(P\), the non-interest-rate factors \(\gamma_t\) affecting prepayment follow a Brownian motion with drift \(\mu_\gamma\) and volatility \(\sigma_\gamma\)

\[
d\gamma_t = \mu_\gamma dt + \sigma_\gamma dZ_{\gamma t},
\]

and that the arrival rate of the Poisson liquidity jump is \(\mu_t\). Thus,

\[
\left( E_t^Q - E_t^P \right) \left[ \partial P_{k,jt} \over \partial \gamma_t \right] d\gamma_t = \frac{\partial P_{k,jt} \over \partial \gamma_t \sigma_\gamma \mu_t}{\partial \gamma_t} + \frac{1}{dt} \left( E_t^Q - E_t^P \right) [dJ_t] = \mu_t \left( \pi_{\mu t} - 1 \right) .
\]

Substituting into (E.3) – (E.4), we obtain that the OAS is given by (3.2) and the \(OAS^Q\) by (3.5).
F Pricing model details

F.1 Interest rate model

We assume that swap rates follow a three-factor Heath, Jarrow, and Morton (1992) (HJM) model. Let \( f(t, T) \) denote the time \( t \) instantaneous forward interest rate for risk-free borrowing and lending at time \( T \). We model the forward rate dynamics under the (interest rate) risk-neutral measure as

\[
df(t, T) = \mu_f(t, T) \, dt + \sum_{i=1}^{3} \sigma_{f,i}(t, T) \, dW^Q_{it},
\]

where \( W^Q_{it} \) are independent standard Weiner processes under the risk-neutral measure \( Q \), and, under no arbitrage, the expected change in the forward rate is given by

\[
\mu_f(t, T) = \sum_{i=1}^{3} \sigma_{f,i}(t, T) \int_{t}^{T} \sigma_{f,i}(t, u) \, du.
\]

Thus, the risk-neutral dynamics of the instantaneous forward rate are completely determined by the initial forward rate curve and the forward rate volatility functions, \( \sigma_{f,i}(t, T) \). Similarly to Trolle and Schwartz (2009), we assume that the volatility function of each factor \( \sigma_{f,i}(t, T) \)

\[
\sigma_{f,i}(t, T) = (\alpha_{0,i} + \alpha_{1,i}(T-t)) e^{-\gamma_i(T-t)}. \tag{F.1}
\]

This specification has the advantage of allowing for a wide range of shocks to the forward rate curve while ensuring that the forward rate model above is Markovian.

Trolle and Schwartz (2009) show that, setting the volatility of the forward rates to be as in (F.1), the time \( t \) price of a zero-coupon bond maturing at time \( T \), \( P(t, T) \), is given by

\[
P(t, T) = \exp \left\{ - \int_{t}^{T} f(t, u) \, du \right\} = \frac{P(0, T)}{P(0, t)} \exp \left\{ \sum_{i=1}^{3} B_{x,i}(T-t) x_{it} + \sum_{i=1}^{3} \sum_{j=1}^{6} B_{\phi,i}(T-t) \phi_{ji,t} \right\},
\]

where the state variables \( \{x_{it}, \phi_{ji,t}\} \) follow

\[
dx_{it} = -\gamma_i x_{it} \, dt + dW^Q_{it} \\
d\phi_{1,i,t} = (x_{it} - \gamma_i \phi_{1,i,t}) \, dt \\
d\phi_{2,i,t} = (1 - \gamma_i \phi_{2,i,t}) \, dt \\
d\phi_{3,i,t} = (1 - 2\gamma_i \phi_{3,i,t}) \, dt \\
d\phi_{4,i,t} = (\phi_{2,i,t} - \gamma_i \phi_{4,i,t}) \, dt \\
d\phi_{5,i,t} = (\phi_{3,i,t} - 2\gamma_i \phi_{5,i,t}) \, dt \\
d\phi_{6,i,t} = (2\phi_{5,i,t} - 2\gamma_i \phi_{6,i,t}) \, dt.
\]
The coefficients \( \{ B_{x_i}, B_{\phi_j} \} \) are functions of the parameters of the volatility function and the time to maturity \( \tau = T - t \), and are given by

\[
B_{x_i} (\tau) = \frac{\alpha_{1i}}{\gamma_i} \left( \left( \frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}} \right) (e^{-\gamma_i \tau} - 1) + \tau e^{-\gamma_i \tau} \right)
\]

\[
B_{\phi_i} (\tau) = \frac{\alpha_{1i}}{\gamma_i} (e^{-\gamma_i \tau} - 1)
\]

\[
B_{\phi_{2i}} (\tau) = \left( \frac{\alpha_{1i}}{\gamma_i} \right)^2 \left( \frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}} \right) \left( \frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}} \right) (e^{-\gamma_i \tau} - 1) + \tau e^{-\gamma_i \tau}
\]

\[
B_{\phi_{3i}} (\tau) = -\frac{\alpha_{1i}}{\gamma_i^2} \left( \frac{\alpha_{1i}^2}{\gamma_i^2} + \frac{\alpha_{0i}^2}{\gamma_i} + \frac{2 \alpha_{0i} \alpha_{1i}}{\gamma_i} \right) (e^{-2\gamma_i \tau} - 1) + \left( \frac{\alpha_{1i}}{\gamma_i} + \frac{\alpha_{0i}}{\gamma_i} \right) \tau e^{-2\gamma_i \tau} + \frac{\alpha_{1i}}{2} \tau^2 e^{-2\gamma_i \tau}
\]

\[
B_{\phi_{4i}} (\tau) = -\frac{1}{2} \left( \frac{\alpha_{1i}}{\gamma_i} \right)^2 (e^{-2\gamma_i \tau} - 1).
\]

Consider now a period of length \( \nu \) and a set of dates \( T_j = t + \nu j, j = 1, \ldots, n \). The time \( t \) swap rate for the period \( t \) to \( T_n \), with fixed-leg payments at dates \( T_1, \ldots, T_n \) is given by

\[
S (t, T_n) = \frac{1 - P (t, T_n)}{\nu \sum_{j=1}^{n} P (t, T_j)}, \quad (F.2)
\]

and the time \( t \) forward swap rate for the period \( T_m \) to \( T_n \), and fixed-leg payments at dates \( T_{m+1}, \ldots, T_n \) by

\[
S (t, T_n) = \frac{P (t, T_m) - P (t, T_n)}{\nu \sum_{j=m+1}^{n} P (t, T_j)}.
\]

Applying Ito’s lemma to the time \( u \) forward swap rate between \( T_m \) and \( T_n \), and switching to the forward measure \( Q^{T_m, T_n} \) under which forward swap rates are martingales (see e.g. Jamshidian, 1997), we obtain

\[
dS (u, T_m, T_n) = \sum_{i=1}^{3} \sum_{j=m}^{n} \zeta_j (u) B_{x_i} (T_j - u) dW_{iu}^{Q^{T_m, T_n}},
\]
where

\[ \zeta_j(u) = \begin{cases} 
\frac{P(u, T_m)}{\nu \sum_{j=m+1}^{n} P(u, T_j)} & \text{if } j = m; \\
\frac{P(u, T_m)}{\nu \sum_{j=m+1}^{n} P(u, T_j)} - v S(u, T_m, T_n) \frac{P(u, T_j)}{\nu \sum_{j=m+1}^{n} P(u, T_j)} & \text{if } j = m + 1, \ldots n - 1 \\
(1 + v S(u, T_m, T_n)) \frac{P(u, T_n)}{\nu \sum_{j=m+1}^{n} P(u, T_j)} & \text{if } j = n.
\end{cases} \]

Notice that, since the \( \zeta_j(u) \) terms are stochastic, the forward swap rates are not normally distributed. We can, however, approximate \( \zeta_j(u) \) by their time \( t \) expected values, which are their time \( t \) values since these terms are martingales under the forward-swap measure. Thus, given date \( t \) information, the swap rate between dates \( T_m \) and \( T_n \) is (approximately) normally distributed

\[ S(T_m, T_n) \sim N(S(t, T_m, T_n), \sigma_N(t, T_m, T_n) \sqrt{T_m - t}) \]

where the volatility \( \sigma_N \) is given by

\[ \sigma_N(t, T_m, T_n) = \left( \frac{1}{T_m - t} \int_t^{T_m} \sum_{i=1}^{N} \left( \sum_{j=m}^{n} \zeta_j(t) B_{j_i}(T_j - u) \right)^2 du \right)^{\frac{1}{2}}. \]

F.2 Yield curve model

We closely follow the estimation of Gürkaynak et al. (2007) on Treasury yields using quotes on par swap yields with maturities between 1 and 40 years. We assume that instantaneous forward rates \( n \)-years hence are a function of six parameters:

\[ f_t(n, 0) = \beta_0 + \beta_1 \exp(-n/\tau_1) + \beta_2 (n/\tau_1) \exp(-n/\tau_1) + \beta_3 (n/\tau_2) \exp(-n/\tau_2). \]  

We fit these parameters at month end by minimizing the sum of squared deviations between actual and predicted swap prices weighted by their inverse duration, which is approximately equal to minimizing the sum of squared yield deviations.

F.3 Prepayment model

As described in the main text, we begin by constructing a panel of monthly dealer prepayment forecasts by coupon-vintage using data from eMBS and Bloomberg LP. Specifically, we match pool characteristics from eMBS (WAC, WALA, WAM) to corresponding prepayment forecasts from Bloomberg. For each coupon until May 2003, and for each coupon-vintage from May 2003 onward, dealers report a prepayment forecast for each of the nine interest rate scenarios, as well as a WAC and WAM. To obtain additional pool characteristics, for the later sample, each survey is matched to its corresponding pool in eMBS. For the earlier sample, we match the survey to the vintage of the same coupon in eMBS with the minimum Mahalanobis distance based on WAC and WAM.
from the dealer’s response. We only use securities that have a remaining principal balance in eMBS of more than $1 million.

Dealers update their forecasts on different dates, so we use the most recent response as of the end of the month for each dealer (excluding dealers who did not update their response during that month), keeping only those securities in a month for which at least two dealers responded. Because we are interested in extracting prepayment model parameters that capture, for instance, the expectations of the rate-sensitivity of a security, we match each dealer’s response to the swap rate of the day before that dealer’s survey response was updated.

The prepayment forecasts in Bloomberg are reported in “PSA” terms, which can be translated into monthly CPRs using the following formula:

$$\text{CPR}_\tau = \text{PSA}/100 \times \min(0.2 \times \text{Age}_\tau, 6) \text{ for } t \leq \tau \leq \text{WAM}$$  \hspace{1cm} (F.4)

Thus, two securities with the same PSA forecast but of different ages (WALAs) will have different “average” CPRs if at least one of the securities is unseasoned. Because we would like to capture the prepayment speed forecast of the dealers with a single number for ease of estimation, we use the PSA forecast and the WALA\(^2\) to compute the WAL (weighted average life), and thus the WAL-implied long-run CPR, defined as the constant monthly CPR that generates the WAL.

Specifically, we convert the monthly CPRs generated using equation F.4 to SMMs and compute the implied cash flows as in Section 2. The WAL is then defined as:

$$\text{WAL}_t = \frac{\sum_{j=t+1}^{\text{WAM}} \text{CF}_j}{\sum_{j=t}^{\text{WAM}} \text{CF}_j}.$$  \hspace{1cm} (F.5)

This gives us one long-run CPR forecast for each scenario per vintage per dealer. The nine different scenarios give us information about the expected rate sensitivity of the security. A common way to model this rate-sensitivity is through the use of an “S curve” as mentioned in the main text. Such a curve captures the observed behavior that prepayments are low for securities that are “out-of-the-money,” i.e., the incentive to refinance is negative, and are mostly due to turnover and, to a lesser extent, cash-out refinancing or defaults. As a pool moves in-the-money (the refinancing incentive becomes positive) the refinancing component becomes a more important driver of prepayments, but at a declining rate: there is an incentive region in which prepayments are highly sensitive to changes in the interest rate (typically somewhere in the incentive region of 50-150 basis points) while beyond that, there is little sensitivity to further decreases in the available rate.

We convert our nine long-run CPRs into SMMs and fit the following S curve for each dealer

\(^2\)Since dealers don’t actually report WALAs, we infer the WALA for a particular dealer’s response by subtracting that dealer’s surveyed WAM from the average sum of the WAM and WALA in eMBS.
for a vintage using nonlinear least squares:

\[
SMM^{LR}_{i} = b_1 + b_4 \exp \left( \frac{b_2 + b_3 \cdot \text{INC}_i}{1 + \exp (b_2 + b_3 \text{INC}_i)} \right) \text{ for } i = 1, 2, \ldots, 9
\]  

where \( b_1, b_4 \in [0, 1] \) and \( b_1 + b_4 \leq 1 \) (these constraints ensure that the function is bounded by 0 and 1). Here, INC\(_i\) is defined as the difference between the dealer’s observed swap rate and WAC in scenario \( i \).

Estimating an S curve for each dealer allows us to “average” these dealer responses despite the fact that often the surveys were updated on different days and thus refer to slightly different interest rate scenarios. We take this average by averaging fitted dealer SMMs at 50 basis point intervals between -300 and 300 basis points, with the 0-scenario corresponding to the average 0-scenario across dealers.

Finally, because cash flows, and thus the OAS, depend on not just the average long-run prepayment rate, but also the time pattern of prepayments, we fit a series of monthly SMMs in the form of equations (4.1) and (4.2) to the dealer-averaged long-run CPR forecasts. As discussed in the main text, this functional form creates the “burnout effect” of prepayments. However, because the Bloomberg data provide no additional information as to the time pattern of prepayments, it is impossible to jointly identify \( \chi_t, \kappa_1, \kappa_2 \) for each security. We therefore assume that \( \kappa_1 \) and \( \kappa_2 \) are universal parameters and let \( \chi_t \) vary across securities and time. To calibrate \( \kappa_1 \) and \( \kappa_2 \), we exploit the fact that as \( \text{INC} \to \infty \), \( SMM \to b_1 + \chi_2 + (1 - \chi) \kappa_2 \) (for \( \text{WALA} > 30 \)). Thus, \( b_1 + \kappa_1 \) and \( b_1 + \kappa_2 \) represent the speeds that a seasoned pool would prepay at if it were deeply in the money and composed of only fast or only slow borrowers, respectively. We therefore estimate \( \hat{\kappa}_1 = \kappa_1 \) and \( \hat{\kappa}_2 = \kappa_2 \) by taking the 99th and 1st percentiles of survey SMMs (less an average \( b_1 \), which is negligible) for the -300 basis point interest rate scenario among seasoned ITM securities in our sample. This yields \( \hat{\kappa}_1 = 0.11 \) and \( \hat{\kappa}_2 = 0.014. \)

Given \( \kappa_1 \) and \( \kappa_2 \), there are then four coefficients to be estimated for each security on each date: \( \chi_t, b_1, b_2, \text{ and } b_3 \). We fit these four coefficients using nonlinear least squares with the thirteen dealer-averaged long-run fitted CPRs. Because of its flexibility, this model is able to fit the long-run CPRs quite well; the MAE across securities is less than 0.2.

Figure A-2 illustrates that the prepayment model that results from the previous procedure fits subsequently realized prepayment behavior well. We first divide the trusts in our sample into three groups based on their moneyness as of month \( t \): OTM (moneyness<0; 1199 observations); weakly ITM (moneyness \( \in [0, 1) \)); 1573 obs.) and strongly ITM (moneyness>1; 941 obs.). Within each group, we then form deciles based on predicted three-month CPRs, and plot the average realized CPR (over months \( t + 1 \) to \( t + 3 \)) against the average predicted CPR. The figure illustrates a few important points. First, as discussed in Section 4.3, moneyness (the refinance incentive) is the main determinant of predicted and realized CPRs. Second, however, it is not the only determinant: there is overlap between the three groups, meaning that predicted (and realized) CPRs are

\[\text{We have experimented with alternative calibrations, and obtained qualitatively similar results.}\]
not always monotonic in moneyness. This reflects the factors emphasized in our specification in equations (4.1) and (4.2)—seasoning and burnout—as well as variation over time in turnover or refinancing sensitivity. Third, there is a lot less variation in the prepayment speeds of OTM securities than for ITM securities. Last, the model based on the dealer surveys fits actual prepayment behavior well: the dots are aligned along the 45 degree line.

**Figure A-2: Prepayments: Model projections vs. realized.** Figure shows binned scatterplot of realized three-month constant prepayment rate (CPR) over months $t + 1$ to $t + 3$ versus model-predicted three-month CPR at the end of month $t$. Within each of three moneyness groups (defined as of month $t$), ten deciles are formed by predicted CPR, and for each of the resulting 30 groups we plot average predicted and realized CPRs.

### F.4 Monte Carlo simulations

As discussed in Section 2, computing the OAS requires Monte Carlo simulations of swaps and discount rates. Along each simulation, we use the prepayment model to compute MBS cash flows. We take the OAS to be the constant spread to swaps that sets the average discounted value of cash flows along these paths equal to the market price. To construct these paths, we first simulate 1,000 paths of the three factors of the interest rate model using draws of the state variables described in Appendix F.1. We use antithetic variables as a variance reduction technique, giving us 2,000 paths in total.
Identifying \( \text{OAS}^Q \) from the cross-section

In this Appendix, we relax the assumption that the risk-neutral prepayment function is a multiple of the physical prepayment function and explore identification of \( \text{OAS}^Q \) from the cross section of MBS. This alternative identification strategy relies on the cross section of stripped securities being priced fairly relative to each other at any given point in time, without requiring us to make assumptions about the physical prepayment function.

In particular, we model the date \( \tau \) risk-neutral single-monthly mortality (Q-SMM) rate of security \( j \) as

\[
s_{j,\tau}^Q = b_{1Q} \min \left( \frac{\text{Age}_j}{30}, 1 \right) + \kappa^Q \cdot f^j_{\tau} \cdot \frac{\exp \left( b_{2Q} + b_{3Q} \cdot \text{INC}_j^\tau \right)}{1 + \exp \left( b_{2Q} + b_{3Q} \cdot \text{INC}_j^\tau \right)},
\]

where \( \text{INC}_j^\tau \) is the interest rate incentive to refinance and \( f^j_{\tau} \) is the “factor” (a pool’s remaining balance relative to origination), at date \( \tau \). In this prepayment function we use the factor as a proxy for the “burnout” effect; this is a simplified version of the specification in the main text. Mortgage rates have mostly trended lower in our sample and, absent the large volume of data from prepayment surveys, one cannot separately identify empirically the effect of “burnout” (remaining borrowers are largely rate insensitive) from the effect of a lower sensitivity of all borrowers to rates (that would lead to the upper-flatness in the “S-shaped” prepayment function). In the simplified version we assume that securities that exhibited more prepayments in the past (low \( f^j_{\tau} \)) tend to prepay more slowly, implying fewer parameters to be estimated than in the main text (\( \chi \) and \( \kappa \) in equations (4.1) and (4.2) in Section 4.3).

For each pool \( j \in J_\tau \) that trades at date \( \tau \), we compute the OAS implied by a set of parameters \( (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q) \) for the IO and the PO strip on that pool, denoted, respectively, by \( \text{OAS}_{j,\text{IO}} (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q) \) and \( \text{OAS}_{j,\text{PO}} (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q) \). At each date \( \tau \), we then minimize the remaining-principal-balance-weighted sum of the squared difference between OAS on the IO and the PO strips written on the same pool:

\[
\arg \min_{j \in J_\tau} \sum_{j \in J_\tau} \frac{\theta_{j\tau}}{\sum_{j \in J_\tau} \theta_{j\tau}} \left( \text{OAS}_{j,\text{IO}} (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q) - \text{OAS}_{j,\text{PO}} (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q) \right)^2,
\]

where \( \theta_{j\tau} \) is the remaining principal balance in pool \( j \) at date \( \tau \). The risk-neutral OAS on the pass-through security on pool \( j \) (the combination, denoted “PT,” of IO and PO) at date \( \tau \) is then computed as the OAS implied by the set of parameters \( (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q)^* \) that solves equation (G.1) at date \( \tau \):

\[
\text{OAS}_{j,\text{PT}}^Q \equiv \text{OAS}_{j,\text{PT}} ( (b_{1Q}, b_{2Q}, b_{3Q}, \kappa^Q)^* ).
\]

This identification strategy relies on the assumption that the cross section of securities at a
given date are priced fairly relative to each other, thus relaxing both the assumption that the risk-neutral SMM rate is a (pool- and date-specific) multiple of the physical SMM rate and that the IO and the PO on a given pool are equally exposed to non-interest and non-prepayment shocks. In addition, this strategy does not rely on any \( \mathbb{P} \)-measure prepayment function.

Figure A-3 shows the cross-sectional scatterplot of the resulting \( \text{OAS}^Q \) on the passthrough security, together with the local smoother. Though the level of the \( \text{OAS}^Q \) that is identified from the cross-section is somewhat higher than in the estimation that relies on both pricing and physical prepayment information, the overall conclusion remains the same: the \( \text{OAS}^Q \) does not vary significantly with moneyness, suggesting that differences in pool-level liquidity do not drive the OAS smile. This is confirmed in Table A-3, which shows the equivalent regressions to those in Table 4 in the main paper. As was the case there, the only bin that exhibits significantly higher \( \text{OAS}^Q \) than for ATM pools is the deeply ITM bin (moneyness \( \geq 2.25 \)). The coefficient on that bin becomes larger if we add month fixed effects or if we end the sample in July 2007. However, the magnitude remains quite similar to the one in Table 4.

Despite these similarities, this identification strategy has a number of limitations relative to the one used in the main text. The method in this appendix only recovers the risk-neutral prepayment speed but does not allow the estimation of the prepayment risk premium (\( \text{OAS} - \text{OAS}^Q \)), which is central to our analysis. One can only back out the prepayment risk premium from physical prepayment speeds. In our main analysis, we are fitting a parsimonious functional form to market participants’ forecasts of future prepayment speeds under different rate scenarios; this helps against misspecification of the prepayment function relative to the real-time \( \mathbb{P} \)-expectations of market participants. The fact that the “standard” OAS (under \( \mathbb{P} \)-prepayments) look similar as in the dealer data further assuages fears that misspecification is driving our findings about the OAS smile.

**H Zero-volatility spreads**

Because of the prepayment option embedded in MBS, which exposes an MBS investor to interest rate risk, the paper studies mortgage spreads adjusted for the prepayment option, or OAS. As discussed, the value of the prepayment option depends not only on the expected path of interest rates, but also on their volatility; the OAS measure correctly takes this into account. In this section we briefly discuss an alternative spread measure considered by market participants, the zero-volatility spread (ZVS, also called Z-spreads).

Denoting with \( P_M \) the market price of an MBS, the ZVS is defined by:

\[
P_M = \sum_{k=1}^{T} \frac{X_k(\mathbb{E}r_k)}{\prod_{j=1}^{k}(1 + \mathbb{E}r_j + \text{ZVS})}. \tag{H.1}
\]

Differently from the OAS, whose formula appears in (2.1), in computing the ZVS both cash
Figure A-3: Cross-sectional variation in OAS\(^Q\), based on alternative identification from prices only. This figure shows the scatterplot and local smoother of the cross-sectional variation in OAS\(^Q\) on pass-through securities, estimated using the cross section of IO and PO prices each month.
flows and discounts are evaluated along a single expected risk-neutral rate path, thus ignoring the effects of uncertainty about the timing of prepayments on the MBS valuation. This implies that the ZVS will be larger than the OAS.\(^4\) Because ZVS abstract from rate uncertainty, following practitioners (for example, Hayre, 2001) we refer to the ZVS-OAS difference as the “option cost”:\(^5\)

\[
\text{Option cost} \equiv \text{ZVS} - \text{OAS}. \quad (H.2)
\]

In Figure A-4, we compare the time series of the ZVS on the TBA index to the OAS.\(^6\) The ZVS (gray line) is typically around 50 to 100 basis points, but, similar to the OAS (black line), the ZVS spiked in the fall of 2008. The ZVS also reached high levels around the 1998 LTCM turmoil, and in 2002 and 2003 in conjunction with the unprecedented refinancing wave in mortgage markets. In Table A-4, we regress ZVS on four risk measures that we considered in Section 2.2. As for the OAS, we find the ZVS to be also related to the Baa-Aaa spread. Unlike for the OAS, we find that the average ZVS is strongly related to implied volatility. Intuitively, as with other American options, the value of the prepayment option increases in the volatility of the underlying.

Figure A-5 documents the cross-sectional variation in the ZVS and option cost. As shown in

---

\(^4\) Market participants also consider simple yield spreads (YS). As for the ZVS, the YS computes cash flows abstracting from uncertainty but discounts at \(((1 + \text{YS} + y^*)\), where \(y^*\) is the swap rate with a duration that is closest (or equal by interpolating rates for the two rates with the closest duration) to the MBS duration. Unlike the ZVS, the YS discounts all cash flows at a constant rate \(y^*\), which is independent of cash flow timing, and as a result, changes in the slope of the yield curve (or duration of the MBS) will not be reflected in the YS.

\(^5\) We note, however, that this is a slight misnomer, as adding volatility to discount rates (in the denominator of equation (H.1)) increases the model value and therefore raises OAS relative to ZVS. In practice, this countervailing effect is small, such that ZVS almost always exceed OAS.

\(^6\) Both time series and cross-sectional patterns of ZVS and option costs coming from our pricing models have similar properties to those shown here.
Panel (b), ZVS are generally increasing in a contract’s moneyness, though the relation flattens out for ITM securities. Panel (c) shows that the option cost (the difference between ZVS and OAS) is hump-shaped, with securities that are closest to par having the largest option cost. The hump shape of the option cost can be understood by analogy to the Vega (sensitivity to changes in the volatility of the underlying) of vanilla call options, which is small for options that are deeply ITM or OTM but large for options near the money. Option execution in MBS are driven by an S-shaped prepayment function (rather than an exercise boundary) as discussed in more detail in Section 4.3 but the patterns are analogous. As shown below, this pattern is directly related to the well-known “negative convexity” (i.e., concavity) of MBS prices with respect to rates, shown in panel (d): as rates drop (the security’s moneyness increases), prices increase less than linearly, especially for near-the-money securities.

We conclude this section by briefly discussing an extension of the model from Section 3 that features interest rate uncertainty to discuss the cross-sectional variation in the option cost. As discussed in Section 2, the option cost is reflected in the difference between the ZVS and the OAS, which with interest rate uncertainty can be shown to equal

$$ZVS_t - OAS_t = -\frac{1}{P_t} \frac{\partial^2 P_t}{\partial r_t^2} \sigma^2_r,$$

where $\sigma_r$ is the volatility in the innovation in the interest rate diffusion. According to this expression, the option cost is positive for negatively convex securities and is proportional to the Gamma of the security ($\partial^2 P_t / \partial r_t^2$). Based on standard results in option pricing, Gamma is generally greatest for at-the-money options and diminishes when moving either in or out of the money; furthermore, the option Vega, that is, its sensitivity to volatility, is directly related to its Gamma. This is consistent with the patterns shown in panels (c) and (d) of Figure A-5: option costs are largest for ATM securities and prices are a concave function of moneyness (and therefore $r$).

---

Footnote: Intuitively, prepayments are not sensitive to small changes in interest rates when the prepayment option is deeply ITM or OTM, so that adding volatility to future interest rates matters little for the expected value of the security; the same is not true for ATM securities.
Figure A-4: Time series of spreads on the TBA index. This figure shows time-series variation in option-adjusted (OAS) and zero-volatility (ZVS) spreads (to swaps) on a value-weighted index based on TBA quotes from six dealers. The data do not contain ZVS prior to 1998. Additional detail is available in Section 2.

Table A-4: Time-series regressions on TBA index. Coefficient estimates from OLS regression of spreads reported in the top row on the Aaa-Treasury spread, the Baa-Aaa spread, the 10-year to 3-month slope of the Treasury curve and the 2-year into 10-year swaption implied volatility. All regressors are standardized to have zero mean and unit standard deviation. Newey-West standard errors (6 lags) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. 

<p>| (1) | (2) | (3) | (4) |</p>
<table>
<thead>
<tr>
<th>OAS</th>
<th>ZVS</th>
<th>OAS</th>
<th>ZVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa - Treas</td>
<td>3.2** [1.5]</td>
<td>6.5*** [2.3]</td>
<td>2.3* [1.2]</td>
</tr>
<tr>
<td>Treas Slope</td>
<td>-0.5 [2.4]</td>
<td>-0.2 [1.3]</td>
<td>2.0 [1.3]</td>
</tr>
<tr>
<td>Swaption Vol.</td>
<td>2.7 [4.2]</td>
<td>15.5*** [2.5]</td>
<td>-0.8 [2.4]</td>
</tr>
<tr>
<td>Const</td>
<td>8.8*** [1.8]</td>
<td>70.9*** [1.8]</td>
<td>3.6*** [1.3]</td>
</tr>
<tr>
<td>Adj. R2</td>
<td>0.76</td>
<td>0.88</td>
<td>0.31</td>
</tr>
<tr>
<td>Obs.</td>
<td>180</td>
<td>156</td>
<td>139</td>
</tr>
<tr>
<td>Dates</td>
<td>199601.201012</td>
<td>199801.201012</td>
<td>199601.200707</td>
</tr>
</tbody>
</table>
Figure A-5: Cross-sectional variation in spreads and prices of MBS in TBA market The panels show scatterplots and local smoothers of the cross-sectional variation in OAS, ZVS, option cost (ZVS-OAS) and price for MBS coupons with remaining principal balance (in 2009 dollars) of $100 million or more.