Federal Reserve Bank of New York
Staff Reports

Understanding Mortgage Spreads

Nina Boyarchenko
Andreas Fuster
David O. Lucca

Staff Report No. 674
May 2014
Revised April 2015

FRBNY Staff REPORTS

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 2015
JEL classification: G10, G12, G13

Abstract

Most mortgages in the United States are securitized in agency mortgage-backed securities (MBS), and thus the yield spreads on these securities are a key determinant of homeowners' funding costs. We study the variation in these spreads, over time and across securities, and document that they display a cross-sectional smile pattern with respect to the securities' coupon rates. We propose non-interest-rate prepayment risk as a candidate driver of the spread variation and present a new pricing model that uses "stripped" MBS prices to identify the contribution of this risk. The pricing model finds the smile to 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 then study the MBS market's response to the Fed's large-scale asset purchases and use the pricing model to interpret the post-announcement divergence of spreads across MBS.

Key words: agency mortgage-backed securities, option-adjusted spreads, prepayment risk, OAS smile

---

Boyarchenko, Fuster, Lucca: Federal Reserve Bank of New York (e-mail: nina.boyarchenko@ny.frb.org, andreas.fuster@ny.frb.org, david.lucca@ny.frb.org). The authors are grateful to John Campbell, Hui Chen, Benson Durham, Laurie Goodman, Arvind Krishnamurthy, Alex Levin, Francis Longstaff, Emanuel Moench, Taylor Nadauld, Stijn Van Nieuwerburgh, Annette Vissing-Jørgensen, Jonathan Wright, and seminar audiences at the Federal Reserve Bank of San Francisco, Morgan Stanley, NYU Stern School of Business, Rutgers University, UCLA Anderson School of Management, and the Toulouse School of Economics for helpful comments and discussions. Karen Shen provided outstanding research assistance.

1 Introduction

At the peak of the financial crisis in the fall of 2008, spreads on residential mortgage-backed securities (MBS) guaranteed by US 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] Spreads on MBS are a key determinant of the cost of residential credit to households since most mortgages in the US are securitized in these securities.[2] Despite the economic importance of these spreads, and the fact that MBS are one of the world's most important fixed income asset classes, little academic attention has been devoted to understanding their determinants. This paper fills this gap by characterizing time-series and cross-sectional spread variation, and by presenting a method to disentangle the contribution of different risk factors to this variation.

Credit risk of MBS is limited because these securities are explicitly (for Ginnie Mae) or implicitly (for Fannie Mae and Freddie Mac) guaranteed by the US government. However, MBS investors face other risks, related in particular to the timing of the cash flows. Mortgage borrowers can freely prepay their mortgage balance at any time, and they do so especially when interest rates decline below those on their current loans. This introduces an option-like feature into MBS and exposes investors to additional interest rate risk. Yields on MBS exceed those on Treasuries or plain-vanilla 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. Conceptually, the OAS is equal to a weighted average of future expected returns after hedging for interest rate risk, suggesting that a positive OAS compensates investors for additional sources of risk. We construct an OAS measure based on surveys of investors' prepayment expectations, but also characterize spread variation directly using quotes from six different dealers over a period of 15 years. In the time series, we find that OAS (to swaps) on a market value-weighted index are typically close to zero but reach 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.

[1] http://www.federalreserve.gov/newsevents/press/monetary/20081125b.htm.

[2] 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 totalled about $6 trillion in principal outstanding. There are also "private-label" residential MBS issued by private firms (and backed by subprime, Alt-A, or jumbo loans), as well as commercial MBS.

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. Moneyness is a key distinguishing feature of an MBS as it determines borrowers' incentive to prepay their loans. We uncover an "OAS smile" in the MBS cross section: 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). We find a similar smile pattern in hedged MBS returns. Correspondingly, a pure long strategy in deeply ITM MBS earns a Sharpe ratio of about 1.3 in our sample, as compared to about 0.5 for a long-ATM strategy. We also find that past OAS predict MBS returns, consistent with their interpretation as weighted expected excess returns.

The OAS smile suggests that investors in MBS earn risk compensation for factors other than interest rates, and that MBS have differential exposures to these risks as a function of their moneyness. Aside from interest rates, borrower prepayments are driven by other factors such as house prices, underwriting standards, and government policies. Variability in these non-interest-rate prepayment factors is not easily diversifiable and is thus a key risk factor to MBS investors, and will be reflected in the OAS. While the OAS accounts for the predicted path of 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.[3]

While prepayment risk premia may generate the OAS smile, risk factors unrelated to prepayment, 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. We propose a new approach to extract the total compensation for prepayment risk using market prices of "stripped" MBS, securities that pass through only interest payments (an "IO" strip) or principal payments (a "PO" strip). We show that the additional information provided by separate prices for the interest and principal payments on a given MBS, and the assumption that a pair of strips is fairly valued relative to each other, is sufficient to identify market-implied risk-neutral ("Q")

[3]Importantly, 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.

prepayment rates as multiples of physical ("$\mathbb{P}$") ones. The prepayment risk premium component in the OAS is then simply the difference between the OAS computed using physical ($OAS^{\mathbb{P}}$) and risk-neutral prepayments ($OAS^{\mathbb{Q}}$).

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 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 $OAS^{\mathbb{Q}}$ component. We show that $OAS^{\mathbb{Q}}$ 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 the liquidity of agency securities. Even after controlling for agency debt spreads, OAS are strongly correlated with credit spreads (Baa-Aaa), which is somewhat surprising given the different sources of risk in the two markets. One interpretation of this finding is the existence of a common marginal investor with time-varying risk aversion. Risk aversion may increase when the intermediary between the markets becomes constrained in his risk-bearing capacity (for example, Shleifer and Vishny, 1997; Duffie, 2010; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014), leading to increased risk premia in both markets. Consistent with limited risk-bearing capacity, we find that the supply of MBS, measured as issuance relative to mark-to-market equity of brokers and dealers, is also a significant determinant of $OAS^{\mathbb{Q}}$.

The response of OAS to the Fed's large-scale asset purchases (LSAPs) provides further evidence on the potential role of balance sheet capacity of financial intermediaries in the MBS market. Following the Fed announcement on November 25, 2008, the $OAS^{\mathbb{Q}}$ narrowed across coupons as the Fed reduced the outstanding stock of MBS available to private investors, relaxing their balance sheet constraints. While the $OAS^{\mathbb{Q}}$ declined across different securities, the standard OAS displayed a heterogeneous response: spreads diverged with lower coupon OAS falling and spreads on higher coupons moderately increasing. According to our pricing model, the decline in $OAS^{\mathbb{Q}}$ for higher-coupon MBS is offset by an increase in prepayment risk premia as these securities moved further in the money. This effect is thus a manifestation of the smile pattern in the prepayment risk premium component in the OAS. This channel is distinct from other risk premium channels of LSAPs that have been emphasized in prior studies (for example, Gagnon et al., 2011; Greenwood and Vayanos, 2014; Krishnamurthy and Vissing-Jorgensen, 2013).

A number of papers have studied the interaction of interest rate risk in MBS and other markets. This work 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., 2013; 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 to account for the component of MBS prices not explained 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. Song and Zhu (2014) examine determinants of implied financing rates in mortgage dollar rolls, including MBS supply.

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 consider liquidity as an additional risk factor in the OAS and also show that the OAS smile is in fact a result of prepayment risk but of a more general form than what Gabaix et al. assume. 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 instead relaxes this assumption. Furthermore, we provide a characterization of spread patterns over a long sample period, present a conceptual framework to rationalize our findings, and study risk premia covariates.

From a methodological perspective, this paper is related to credit risk studies that confront their models with both physical and risk-neutral (that is, pricing) data to evaluate the role of default risk premia. Driessen (2005), for example, uses US bond price data and historical default rates to estimate a default event risk premium. Driessen 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

[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.

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.

2 Facts about mortgage spreads

In this section we provide a brief overview of MBS and define the OAS. We then characterize the time-series and cross-sectional spread variation in terms of a few stylized facts. We finally relate the OAS to future expected returns hedged for interest rate risk, and then study this relation in the data.

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. However, while Ginnie Mae securities have the full faith and credit of the US federal 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 sce-

[5] In addition to these "lender swap" transactions, Fannie Mae and Freddie Mac also conduct "whole loan conduit" transactions, where they 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).

nario, as was the case 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 in the timing of future cash flows, X_t. Prepayment rates depend on loan characteristics as well as macroeconomic factors such as house prices. However, variation in interest rates r_t is the most important source of uncertainty affecting prepayments and thus the value of the embedded option. We highlight this fact by denoting cash-flows by $X_t(r_t)$.[8] The model price (which may differ from the market price) of an MBS is equal to the average discounted value of possible cash-flow paths $(X_t(r_t))$. The option-adjusted spread (OAS) is the constant spread to baseline rates that equates model and market prices:

$$P_t = \mathbb{E}_t^* \left[\sum_{k=1}^{T-t} \frac{X_{t+k}\left(r_{t+k}\right)}{\prod_{j=1}^{k}\left(1 + r_{t+j} + \text{OAS}\right)} \right], \tag{2.1}$$

where P_t the market price of an MBS, and $\mathbb{E}^*$ is the expectation under the interest-rate-risk-neutral measure $\mathbb{Q}^*$. The OAS increases the greater the discounted 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.[9]

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

[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.

[8] Appendix A provides a detailed description of how MBS cash flows depend on prepayments and scheduled amortization.

[9] In Appendix F, 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.

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. 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). We consider 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.[10] 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.[11] 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 and data-quality filters are available in Appendix B.

2.2 OAS variation in the time series and cross section

At each point in time, MBS with different coupons coexist. These coupon differences primarily reflect disparate loan rates of the mortgages underlying the MBS, which arise due to variation in loan origination dates, as well as other factors such as "points" paid (or received) by the borrowers.[12]

[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] 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

The benchmark contract in the TBA market is the so-called current coupon, which is a synthetic 30-year fixed-rate MBS obtained by interpolating the highest coupon below par and the lowest coupon above par.[13] The interest in this benchmark is due to the fact that most newly originated mortgages are securitized in coupons trading close to par, so that the current coupon rate can be thought of as the relevant secondary market rate for borrowers seeking a new loan.

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 relatively small 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 percent accounted for only about 20 percent of the total outstanding on a market value-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 generally close to zero, consistent with the limited credit risk of MBS; 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 briefly study the relation between the OAS on the TBA index and 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

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.

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 that there are common pricing factors between the MBS and corporate bond markets.[15] In contrast, implied rate volatility does not explain the OAS variation, consistent with the prediction that since the OAS adjusts for interest rate risk, it 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.

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

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

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 mortgage payment—the borrower's option is "out-of-the-money" (OTM). Aside from determining the refinancing propensity of a loan,

[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.

[15] Brown (1999) relates the OAS to Treasuries of pass-through MBS over the period 1993–1997 to other risk premia 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.

moneyness also measures an investor's gains or losses (in terms of coupon payments) if a mortgage underlying the security prepays (at par) and he reinvests the proceeds in a "typical" newly originated MBS (which will approximately have a coupon equal to the FRM rate at time 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 on average exceed those on ATM securities by 50 basis points or more. Table 2 shows that the pattern shown in the bottom panel of Figure 2 is robust to controlling for month fixed effects (meaning that only cross-sectional variation is exploited) and to ending the sample before the onset of the financial crisis. Instead of imposing parametric restrictions, we simply regress OAS on 50-basis-point moneyness bins, with $[-0.25, 0.25)$ as the omitted category. Comparing columns (1) to (2) and (3) to (4), we see that the addition of month fixed effects matters little for the relative OAS in the cross section. Additionally, we find that the relative spreads of OTM coupons are robust to removing the financial crisis period from the sample, while for ITM coupons, the spread difference to ATM securities was somewhat smaller pre-crisis, but still highly statistically significant.

2.3 OAS as a measure of expected hedged returns

While the OAS is a valuation measure that is widely tracked by financial market participants, it has been called into question because of its model dependence. For instance, Kupiec and Kah (1999) argue that the OAS is the result of misspecification of the prepayment model relative to market expectations, and does not carry noteworthy asset pricing significance. In this section, we address the relevance of the OAS as a measure of risk premia by studying its relation to expected MBS returns. Conceptually, we show that the OAS is a weighted average of expected future excess return on an MBS when interest rate risk is hedged. Consistent with this interpretation, Breeden (1994) provides evidence that OAS do predict subsequent hedged returns over 1988-94. We extend Breeden's results to our sample and also show that, as for the OAS, these returns display a cross-sectional smile pattern.

Formally, let rx_{t+1} be the one-period MBS excess return:

$$1 + r_{t+1} + rx_{t+1} = \frac{P_{t+1} + X_{t+1}\left(r_{t+1}\right)}{P_t},$$

where r_{t+1} is the one-period risk-free rate. Taking expectations of both sides of this expression under the interest-rate-risk-neutral measure $\mathbb{Q}^*$, we obtain

$$1 + r_{t+1} + \mathbb{E}_t^* \left[rx_{t+1} \right] = \frac{\mathbb{E}_t^* \left[P_{t+1} + X_{t+1} \left(r_{t+1} \right) \right]}{P_t},$$

or, equivalently,

$$P_t = \mathbb{E}_t^* \left[\frac{P_{t+1} + X_{t+1} \left(r_{t+1} \right)}{1 + r_{t+1} + \mathbb{E}_t^* \left[rx_{t+1} \right]} \right].$$

Iterating this expression forward, we can thus express the time t price of the security as

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

Comparing this to (2.1), the OAS is related to the path of the one-period expected excess returns under $\mathbb{Q}^*$. Approximating (2.2) and (2.1) around a zero spread and excess return and solving for OAS, we get (see Appendix C):

$$OAS_t \approx \mathbb{E}_t^* \left[\sum_{k=1}^{T-t} w_{kt} \mathbb{E}_{t+k-1}^* \left[rx_{t+k} \right] \right], \tag{2.3}$$

where the weights w_{kt} are given by

$$w_{kt} = \sum_{l=k}^{T-t} \frac{\frac{X_{t+l}(r_{t+l})}{\prod_{j=1}^{l}(1+r_{t+l})} \frac{1}{1+r_{t+k}}}{\mathbb{E}_t^* \left[\sum_{k=1}^{T-t} \frac{1}{1+r_{t+k}} \sum_{l=k}^{T-t} \frac{X_{t+l}(r_{t+l})}{\prod_{j=1}^{l} \left(1+r_{t+j} \right)} \right]}.$$

Thus, 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 $\mathbb{Q}^*$, with the weight declining in horizon.

The construction of realized MBS returns is complicated due to the existence of a large number of securities, different MBS pricing conventions, and security-specific prepayments. In order to study the link between OAS and returns, we therefore rely on return series from the MBS subcomponent of the Barclays Aggregate Bond Index, which is a key benchmark for fixed income index funds.[17] The Barclays data consists of monthly returns by coupon over the same sample pe-

[17]The return on an MBS security 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

riod as in the previous subsection (1996–2010). These returns are available both unhedged (that is, as measured from MBS prices and prepayments alone) and interest-rate-hedged (with a duration-matched swap portfolio). Since the OAS is computed under the $\mathbb{Q}^*$ interest rate measure, we focus on hedged returns (but also analyze unhedged returns for robustness). As we show in Appendix C, expected excess hedged returns under $\mathbb{Q}^*$ coincide with expected excess hedged returns under the physical measure $\mathbb{P}$, which we measure in the data. Finally, while conceptually the OAS corresponds to a weighted average of expected excess returns over the lifetime of the security (as shown in (2.3)), empirically we take advantage of the fact that these weights decline in horizon and only use cumulative returns over the following 12 months, in order to get a sufficiently large number of observations.

Table 3 shows results of a predictive regression of one-year MBS returns on the OAS for the coupons available in the Barclays index (and using the same average OAS across dealers as in the previous subsection). The first column shows the pooled regression, for which we see a positive and highly significant coefficient, and substantial explanatory power.[18] The second column adds month fixed effects, and the coefficient remains strongly significant, meaning that the relation between the OAS and returns also holds in the cross section. The final two columns repeat the same exercise for unhedged returns. In column (3) the coefficient on the OAS is only marginally significant and the explanatory power very low, since most of the variation in unhedged returns is driven by interest rate fluctuations. However, column (4) shows that once aggregate effects are absorbed in the time fixed effect, the OAS again becomes a strongly significant predictor of returns in the cross section. Thus, the OAS predicts future returns as implied by equation (2.3).

Looking directly at the cross-sectional pattern of realized returns, Figure 3 shows that one-year hedged returns across coupons display a smile pattern similar to the one shown in the bottom panel of Figure 2. To assess the economic magnitude, we compute (annualized) Sharpe ratios based on monthly Barclays index returns for portfolios of ITM, ATM and OTM MBS securities over our sample period 1996-2010.[19] Relative to ATM securities, we find much larger Sharpe ratios for OTM and especially ITM securities: using unhedged returns (minus the risk-free rate[20]),

information to compute the paydown return. For additional information see, for example, Lehman Brothers (2008).

[18] Realized returns in the sample move more than 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.

[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] Obtained from Ken French's website.

the Sharpe ratio of a long-ITM (long-OTM) portfolio has a Sharpe ratio of 1.33 (0.74), while the one for a long-ATM position is only 0.54. Looking at hedged returns, we find a very similar pattern, with an ATM Sharpe ratio of essentially zero, whereas the ITM and OTM portfolios obtain Sharpe ratios of 0.74 and 0.25.

In sum, while the OAS is a model-dependent measure, we find it to be empirically related to future realized returns, and document a smile pattern in the cross section of hedged returns similar to that of the OAS. 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.

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 OAS.

Consider a mortgage pool j with coupon rate c_j and remaining principal balance θ_{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 D shows that the main expression of interest is the same if we assume that rates follow a diffusion. Suppose the prepayment rate is given by

$$s_{jt} = f\left(\gamma_t, c_j - r\right).$$

The parameters γ_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 μ_t, the whole market experiences a liquidity event in which a pool j loses a fraction α_{jt} of its market value. Thus α_{jt} should be thought of as how well the security performs in a "bad" market, similar to Acharya and Pedersen (2005). Alternatively, α_{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

$$P_t = \mathbb{E}_t \left[\int_0^{+\infty} \frac{M_{t+s}}{M_t} X_{t+s} ds \right]. \tag{3.1}$$

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 $\mathbb{Q}^*$ 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

$$dR_t \equiv \frac{dP_t}{P_t} + \frac{X_t}{P_t} dt.$$

The no-arbitrage restriction (3.1) implies that the expected return can be represented as

$$\mathbb{E}_t \left[dR_t \right] = r\, dt - \mathbb{E}_t \left[\frac{dM_t}{M_t} \frac{dP_t}{P_t} \right],$$

where r is the risk-free rate. As shown in Section 2.3, 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 π_{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 D that the OAS is given by

$$OAS_{jt} = \alpha_{jt}\mu_t\left(\pi_{lt} - 1\right) - \pi'_{\gamma t}\sigma_\gamma \frac{1}{P_{jt}}\frac{\partial P_{jt}}{\partial \gamma_t}. \tag{3.2}$$

Thus, differences in OAS across securities could be the result of (i) differential exposure α_j to the liquidity shock, or (ii) differential price sensitivity to the prepayment parameters, that is, differential exposure to prepayment risk.

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 ($\partial P/\partial\gamma = \partial P/\partial s \cdot \partial s/\partial\gamma$), we get the following approximate expression for the OAS:

$$OAS_{jt} \approx \alpha_{jt}\mu_t\left(\pi_{lt} - 1\right) + \pi'_{\gamma t}\sigma_\gamma \frac{c_j - r}{\left(s_{jt} + c_j\right)\left(s_{jt} + r\right)}\frac{\partial s_{jt}}{\partial \gamma_t}. \tag{3.4}$$

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, $\bar{s}_j$ corresponds to the expected prepayment speed on security j.

Case 1: $s_{jt} = \bar{s}_j + \gamma_{1t}\beta_j$**.** This is essentially the framework studied by Gabaix et al. (2007). Each pool has a constant exposure β_j to a single market-wide prepayment shock γ_1. Under this functional form, the OAS_{jt} in (3.4) will be a linear function of moneyness $c_j - r$ (regardless of the sign of the risk price $\pi_{\gamma_{1t}}$). This case is therefore inconsistent with the OAS smile.

Case 2: $s_{jt} = \bar{s}_j + \gamma_{1t}(c_j - r_t)$**.** 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.[21] This functional form implies that when ITM securities prepay faster than expected (a positive shock to γ_1), OTM securities prepay slower than expected. This may arise be-

[21] $\gamma_1 > 0$ because in practice prepayments are a non-decreasing function of the incentive to refinance.

cause of mortgage originators' capacity constraints in (larger than expected) refinancing waves.[22] It is easy to see from (3.4) that this case would lead the OAS to be quadratic in $c_j - r$ (the risk price $\pi_{\gamma_{1t}}$ is positive in this case, since every security has a positive exposure to γ_1), and therefore could rationalize the OAS smile.

Case 3: $s_{jt} = \bar{s}_j + \gamma_{1t}\mathbf{1}_{c_j<r} + \gamma_{2t}\mathbf{1}_{c_j\geq r}$. In this multi-factor formulation, OTM and ITM prepayments are driven by different shocks (which for simplicity we assume to be orthogonal). For instance, γ_{1t} might represent the pace of housing turnover while γ_{2t} 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 γ_{1t} risk (since OTM securities benefit from fast prepayment) and a positive exposure to γ_{2t} 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.

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).[23] To separate liquidity and prepayment risk premia, in the next section we provide a method to identify a "prepayment-risk-neutral OAS," denoted $OAS^{\mathbb{Q}}$, as the spread that only reflects liquidity risk:

$$OAS^{\mathbb{Q}}_{jt} = \alpha_{jt}\mu_t\left(\pi_{lt} - 1\right).$$

The prepayment risk premium paid to the investor is then just the difference between the OAS and $OAS^{\mathbb{Q}}$.

[22] 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. (2013) show that originators' profit margins are strongly correlated with mortgage application volume, consistent with the presence of capacity constraints.

[23] This 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."

4 Pricing model: Decomposing the OAS

In this section, we propose a method to decompose the "standard" OAS (or $OAS^{\mathbb{P}}$) into a prepayment risk premium component and a remaining risk premium ($OAS^{\mathbb{Q}}$). 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$^{\mathbb{Q}}$

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 ($OAS^{\mathbb{Q}}$) 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 ("$\mathbb{Q}$") prepayment function is a multiple Λ of the physical ("$\mathbb{P}$") one. We allow the multiplier Λ to be pool-specific to account for differences in pools' sensitivities to non-interest rate sources of prepayment risk. Pricing information on a standard pass-through MBS alone is insufficient to identify Λ, because a single observable (the price) can only determine one unknown (the spread) in the pricing model, leaving Λ 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 experience early accrual (sum of grey areas) and therefore the value of the PO strip increases.

We exploit the differential exposure of the two strips to prepayments to identify $OAS^{\mathbb{Q}}$ and Λ, 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 grey) is about zero. As Λ 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 (red line) is also negative, 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 Λ as the crossing of the OAS IO and PO schedules at the point where the residual risk premium ($OAS^{\mathbb{Q}}$) on the two strips is equalized.[24] 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 ($OAS^{\mathbb{P}}$) and at the market-implied one ($OAS^{\mathbb{Q}}$) is then equal to the prepayment risk premium paid on the pass-through. More formally:

Proposition 4.1. *If 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 using the market-implied prepayment speed, the remaining risk premia are equalized on the strips and recombined pass-through, so that*

$$OAS^{\mathbb{Q}}_{IO,j} = OAS^{\mathbb{Q}}_{PO,j} = OAS^{\mathbb{Q}}_{j}.$$

(In Appendix D we show how this proposition applies to the theoretical setting of Section 3.) 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 $OAS^{\mathbb{P}} - OAS^{\mathbb{Q}}$.*

We apply this method to each IO/PO pair in our sample, thereby identifying pool- and date-specific Λ and $OAS^{\mathbb{Q}}$. This allows us to study time-series and cross-sectional variation in the $OAS^{\mathbb{Q}}$ 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.[25]

The key to this identification is the assumption that $OAS^{\mathbb{Q}}$ are equal across IO and PO strips on the same pool. One could relax this assumption by imposing a parametric form linking $OAS^{\mathbb{Q}}$

[24] MBS market participants sometimes calculate "break-even multiples" similar to our Λ but, to our knowledge, do not seem to track them systematically as measures of risk prices.

[25] An alternative way to identify Λ would be to assume that the OAS reflects only prepayment risk. With this approach, Levin and Davidson (2005) obtain a $\mathbb{Q}$ 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.

(or Λ) across pools. Alternatively one could use TBA prices (as a proxy for the value of the recombined pass-through) and make assumptions on the differential liquidity of the stripped and recombined securities to identify Λ. That said, the impact on the prepayment risk premium and $OAS^{\mathbb{Q}}$ on the pass-through will be limited for reasonable liquidity differences between IOs and POs. For example, we find that assuming $OAS^{\mathbb{Q}}_{PO}$ to be 50 basis points higher than $OAS^{\mathbb{Q}}_{IO}$ never changes $OAS^{\mathbb{Q}}$ by more than 5 basis points relative to the baseline specification with $OAS^{\mathbb{Q}}_{IO} = OAS^{\mathbb{Q}}_{PO}$. Intuitively, as shown in Figure 5, the slope of the OAS schedule for the pass-through is less steep in Λ than the slopes of the IO and PO schedules, and thus $OAS^{\mathbb{Q}}_{IO} - OAS^{\mathbb{Q}}_{PO}$ differences will have a limited effect on the recombined pass-through.

4.2 Stripped MBS data

To implement this identification, 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.[26] 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).[27] A potential concern is that the IO/PO strips we have are not

[26]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.

[27]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 E.

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., Section 4 of Fabozzi, 2006), 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 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).[28] Carlin et al. (2014) use these data to study the pricing effects of investors'

[28]Until May 2003, dealers provided a single set of forecasts for each coupon (separately for Fannie Mae, Freddie Mac,

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 τ single-month mortality rate (SMM), which is the fraction of a pool that prepays, 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^j_τ).[29] The functional form of our prepayment model is:

$$\mathrm{s}^j_\tau = \chi^j_\tau \mathrm{s}^j_{1,\tau} + \left(1 - \chi^j_\tau\right) \mathrm{s}^j_{2,\tau} \text{ for } t < \tau \leq t + \text{Maturity}_j \tag{4.1}$$

where

$$\mathrm{s}^j_{i,\tau} = b^j_1 \min\left(\text{Age}^j_\tau/30, 1\right) + \kappa_i \cdot \frac{\exp\left(b^j_2 + b^j_3 \cdot \text{INC}^j_\tau\right)}{1 + \exp\left(b^j_2 + b^j_3 \cdot \text{INC}^j_\tau\right)} \text{ for } i = 1, 2. \tag{4.2}$$

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.[30] 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 χ_τ and $1 - \chi_\tau$ and shares $\mathrm{s}_{1,\tau}$ and $\mathrm{s}_{2,\tau}$. This setup is a simplified version of the heterogeneous refinancing 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

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).

[29] 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%.

[30] 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).

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 κ_i that differs across the two groups: $\kappa_1 > \kappa_2$. Since group 1 prepays faster, χ_τ 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 E.

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 χ).

5 Model results

The pricing model produces the standard OAS measure as well as the $OAS^{\mathbb{Q}}$, 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 $OAS^{\mathbb{Q}}$ 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 $OAS^{\mathbb{Q}}$, 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 $OAS^{\mathbb{Q}}$ 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 ($\mathbb{Q}$) and physical ($\mathbb{P}$) 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 Tables 4 and 5, 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 [31] Table 4 shows that there is little systematic pattern in $OAS^{\mathbb{Q}}$ across bins; results in columns (2) and (4) suggest in fact that ATM coupons may have slightly *higher* $OAS^{\mathbb{Q}}$ 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, Table 5 shows that the (slightly tilted) smile pattern shown in panel (d) 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 or more relative to ATM securities.

In sum, while the prepayment risk premium in the cross section is strongly linked to the moneyness of the securities, we find little evidence that this is also the case for the remaining risk premium ($OAS^{\mathbb{Q}}$), 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

[31] 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.

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.

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.[32] Comparing the OAS in Figure 8 to the corresponding measures in 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 $OAS^{\mathbb{Q}}$ to be small and the two series to tightly co-move, meaning that much of the OAS variation results from changes in $OAS^{\mathbb{Q}}$ (grey line). Thus, although it is an important determinant of the cross-sectional variation in spreads, prepayment risk does not appear to be the dominant driver of the OAS time-series variation. Indeed as shown in Figure 1, the value-weighted 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. [33] 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 $OAS^{\mathbb{Q}}$ widened.

We now turn to the determinants of the time-series variation in OAS, and in particular its two components: $OAS^{\mathbb{Q}}$ and the prepayment risk premium. In Table 6, we regress 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), as well as the average squared moneyness of the MBS universe. 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 $OAS^{\mathbb{Q}}$ 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

[32] We do this by first averaging spreads within coupons (weighting by IO/PO values) and then across coupons (weighting by market values based on TBA prices).

[33] 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.

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 $OAS^{\mathbb{Q}}$ and agency debt spreads increased in the fall of 2008, when Fannie Mae and Freddie Mac were placed in conservatorship by the US 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.[34]

Credit spreads (Baa-Aaa) continue to be significantly related to OAS, mostly through $OAS^{\mathbb{Q}}$ rather than the prepayment risk component. The sensitivity of $OAS^{\mathbb{Q}}$ 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 the supply of MBS, measured by issuance relative to mark-to-market equity of brokers and dealers, also explains $OAS^{\mathbb{Q}}$ time-series variation. 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.

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 ($OAS^{\mathbb{P}} - OAS^{\mathbb{Q}}$) 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

[34] 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 $OAS^{\mathbb{Q}}$ also reaches substantially higher levels in October compared to August 2008, despite the reduction in credit risk to investors.

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.[35] 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 fixed-rate mortgage 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 fanned out (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 $OAS^{\mathbb{Q}}$). On the other hand, movements in mortgage rates associated with such purchases alter securities'

[35] 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.

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.[36] In terms of the MBS supply effect, we discussed above how the $OAS^{\mathbb{Q}}$ component is flat across coupons and declines with a reduction in MBS supply. Consistent with this, we find that the $OAS^{\mathbb{Q}}$ evolves similarly for the 4.5, 5, and 5.5 coupons. For the 6 coupon, $OAS^{\mathbb{Q}}$ increases in November and December, before dropping toward the level of the other coupons in January as actual LSAP purchases begin. The $OAS^{\mathbb{Q}}$ effect thus suggests that the LSAP program lowered non-prepayment risk premia across the coupon stack.[37]

The cross-coupon "homogeneous" $OAS^{\mathbb{Q}}$ 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 grey 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 $OAS^{\mathbb{Q}}$ 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, we find that risk premia associated with non-interest-rate prepayment factors explain the OAS smile, meaning that OAS tend to be lower for ATM MBS than for others. In contrast, in the time series, the model results imply that

[36]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.

[37]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 average OAS is primarily driven by non-prepayment risk factors, which are linked to credit spreads, MBS supply, and spreads on other agency debt. These results suggest that risk-bearing capacity of MBS investors, as well as the liquidity and default risk of agency securities, drive aggregate spread variation and are, therefore, important determinants of homebuyers' funding costs.

We use the model to study the effects of the Fed's MBS purchases, which, along with other unconventional monetary policies, have been highly debated. Under one view, the central bank's asset composition is irrelevant and Fed purchases of long-term securities such as Treasuries and MBS matter only to the extent that they boost high-powered money through the Fed's liabilities, or signal commitment to future short-term interest rate policies (Woodford, 2012). Under an alternative view, the Fed's asset composition directly affects asset prices (Bernanke, 2009). Gagnon et al. (2011), building on Greenwood and Vayanos (2014) and others, 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. Krishnamurthy and Vissing-Jorgensen (2011) instead argue that Fed purchases have more nuanced pricing implications and mainly affect the price of the securities being purchased, rather than having cross-asset effects by lowering the price of duration risk.

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^{Q} 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" episode around the June 19, 2013 FOMC meeting, when rates backed up on investors' fears that the Fed would start reducing its purchases earlier than previously thought. 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.[38] 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.[39] 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, instead, implies a different explanation 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 $OAS^{\mathbb{Q}}$) 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.[40] 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 $OAS^{\mathbb{Q}}$ to Fed purchases in the wake of the crisis, the $OAS^{\mathbb{Q}}$ reacts to changes in the outstanding supply of MBS and credit spreads also 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 US mortgage market

[38]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.

[39]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.

[40]The 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.

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., 2013; 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 US and abroad.

References

ACHARYA, V. V. AND L. H. PEDERSEN (2005): "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77, 375–410.

ADRIAN, T. AND N. BOYARCHENKO (2012): "Intermediary Leverage Cycles and Financial Stability," Staff report no. 567, Federal Reserve Bank of New York.

ALMEIDA, H. AND T. PHILIPPON (2007): "The Risk-Adjusted Cost of Financial Distress," *Journal of Finance*, 62, 2557–2586.

ARCIDIACONO, N., L. CORDELL, A. DAVIDSON, AND A. LEVIN (2013): "Understanding and Measuring Risks in Agency CMOs," Working paper no. 13-8, Federal Reserve Bank of Philadelphia.

BERNANKE, B. S. (2009): "The crisis and the policy response," Speech delivered at the "Stamp Lecture," London School of Economics.

BOUDOUKH, J., R. F. WHITELAW, M. RICHARDSON, AND R. STANTON (1997): "Pricing Mortgage-Backed Securities in a Multifactor Interest Rate Environment: A Multivariate Density Estimation Approach," *Review of Financial Studies*, 10, 405–446.

BREEDEN, D. T. (1994): "Complexities of Hedging Mortgages," *Journal of Fixed Income*, 4, 6–41.

BROWN, D. T. (1999): "The Determinants of Expected Returns on Mortgage-Backed Securities: An Empirical Analysis of Option-Adjusted Spreads," *Journal of Fixed Income*, 9.

BRUNNERMEIER, M. K. AND L. H. PEDERSEN (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201–2238.

BRUNNERMEIER, M. K. AND Y. SANNIKOV (2012): "The I Theory of Money," Working paper, Princeton University.

——— (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104, 379–421.

CAMPBELL, J. Y. (2013): "Mortgage Market Design," *Review of Finance*, 17, 1–33.

CARLIN, B. I., F. A. LONGSTAFF, AND K. MATOBA (2014): "Disagreement and Asset Prices," *Journal of Financial Economics*, 114, 226–238.

CHEYETTE, O. (1996): "Implied Prepayments," *Journal of Portfolio Management*, 23, 107–115.

COHLER, G., M. FELDMAN, AND B. LANCASTER (1997): "Price of Risk Constant (PORC): Going Beyond OAS," *Journal of Fixed Income*, 6, 6–15.

DENG, Y. AND J. M. QUIGLEY (2012): "Woodhead Behavior and the Pricing of Residential Mortgages," IRES Working Paper 2012-025, National University of Singapore.

DRIESSEN, J. (2005): "Is Default Event Risk Priced in Corporate Bonds?" *Review of Financial Studies*, 18, 165–195.

DUARTE, J. (2008): "The Causal Effect of Mortgage Refinancing on Interest Rate Volatility: Empirical Evidence and Theoretical Implications," *Review of Financial Studies*, 21, 1689–1731.

DUFFIE, D. (2010): "Presidential Address: Asset Price Dynamics with Slow-Moving Capital," *Journal of Finance*, 65, 1237–1267.

DUNN, K. B. AND J. J. MCCONNELL (1981a): "A Comparison of Alternative Models for Pricing GNMA Mortgage-Backed Securities," *Journal of Finance*, 36, pp. 471–484.

——— (1981b): "Valuation of GNMA Mortgage-Backed Securities," *Journal of Finance*, 36, pp. 599–616.

FABOZZI, F. J., ed. (2006): *The Handbook of Mortgage-Backed Securities*, New York, NY: McGraw-Hill, 6th ed.

FELDHÜTTER, P. AND D. LANDO (2008): "Decomposing swap spreads," *Journal of Financial Economics*, 88, 375–405.

FUSTER, A., L. GOODMAN, D. LUCCA, L. MADAR, L. MOLLOY, AND P. WILLEN (2013): "The Rising Gap between Primary and Secondary Mortgage Rates," *Federal Reserve Bank of New York Economic Policy Review*, 19, 17–39.

FUSTER, A. AND J. VICKERY (2015): "Securitization and the fixed-rate mortgage," *Review of Financial Studies*, 28, 176–211.

GABAIX, X., A. KRISHNAMURTHY, AND O. VIGNERON (2007): "Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market," *Journal of Finance*, 62, 557–595.

GAGNON, J., M. RASKIN, J. REMACHE, AND B. SACK (2011): "The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases," *International Journal of Central Banking*, 7, 3–43.

GREEN, R. K. AND S. M. WACHTER (2005): "The American mortgage in historical and international context," *Journal of Economic Perspectives*, 93–114.

GREENWOOD, R. AND D. VAYANOS (2014): "Bond supply and excess bond returns," *Review of Financial Studies*, 27, 663–713.

GROMB, D. AND D. VAYANOS (2002): "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics*, 66, 361–407.

GÜRKAYNAK, R. S., B. SACK, AND J. H. WRIGHT (2007): "The US Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54, 2291–2304.

HANCOCK, D. AND W. PASSMORE (2011): "Did the Federal Reserve's MBS purchase program lower mortgage rates?" *Journal of Monetary Economics*, 58, 498–514.

HANSON, S. G. (2014): "Mortgage Convexity," *Journal of Financial Economics*, 113, 270–299.

HAYRE, L. (2001): "A Concise Guide to Mortgage-Backed Securities (MBSs)," in *Salomon Smith Barney Guide to Mortgage-Backed and Asset-Backed Securities*, ed. by L. Hayre, New York, NY: Wiley and Sons, chap. 1, 9–68.

HE, Z. AND A. KRISHNAMURTHY (2013): "Intermediary Asset Pricing," *American Economic Review*, 103, 732–770.

HEATH, D., R. A. JARROW, AND A. MORTON (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60, 77–105.

JAMSHIDIAN, F. (1997): "LIBOR and swap market models and measures," *Finance and Stochastics*, 1, 293–330.

KRISHNAMURTHY, A. (2010): "How Debt Markets Have Malfunctioned in the Crisis," *Journal of Economic Perspectives*, 24, 3–28.

KRISHNAMURTHY, A. AND A. VISSING-JORGENSEN (2011): "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," *Brookings Papers on Economic Activity*, pp. 215–287.

——— (2012): "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 120, 233–267.

——— (2013): "The Ins and Outs of LSAPs," Presented at the Federal Reserve Bank of Kansas City's Jackson Hole Symposium.

KUPIEC, P. H. AND A. KAH (1999): "On the Origin and Interpretation of OAS," *Journal of Fixed Income*, 9, 82–92.

LEHMAN BROTHERS (2008): "A guide to the Lehman Global Family of Fixed Income Indices," Fixed Income Research, February.

LEVIN, A. AND A. DAVIDSON (2005): "Prepayment Risk- and Option-Adjusted Valuation of MBS," *Journal of Portfolio Management*, 31, 73–85.

LONGSTAFF, F. A. (2005): "Borrower Credit and the Valuation of Mortgage-Backed Securities," *Real Estate Economics*, 33, 619–661.

MALKHOZOV, A., P. MUELLER, A. VEDOLIN, AND G. VENTER (2013): "Mortgage Risk and the Yield Curve," Working paper, http://ssrn.com/abstract=2235592.

MCCONNELL, J. J. AND S. A. BUSER (2011): "The Origins and Evolution of the Market for Mortgage-Backed Securities," *Annual Review of Financial Economics*, 3, 173–192.

MERTON, R. C. (1972): "An analytic derivation of the efficient portfolio frontier," *Journal of Financial and Quantitative Analysis*, 7, 1851–1872.

PERLI, R. AND B. SACK (2003): "Does Mortgage Hedging Amplify Movements in Long-Term Interest Rates?" *Journal of Fixed Income*, 13, 7–17.

RICHARD, S. F. AND R. ROLL (1989): "Prepayments on fixed-rate mortgage-backed securities," *Journal of Portfolio Management*, 15, 73–82.

SCHWARTZ, E. S. AND W. N. TOROUS (1989): "Prepayment and the Valuation of Mortgage-Backed Securities," *Journal of Finance*, 44, pp. 375–392.

SHLEIFER, A. AND R. W. VISHNY (1997): "The Limits of Arbitrage," *Journal of Finance*, 52, 35–55.

SONG, Z. AND H. ZHU (2014): "Mortgage Dollar Roll," Working paper, Federal Reserve Board and MIT Sloan.

STANTON, R. (1995): "Rational Prepayment and the Valuation of Mortgage-Backed Securities," *Review of Financial Studies*, 8, 677–708.

STROEBEL, J. AND J. B. TAYLOR (2012): "Estimated Impact of the Federal Reserve's Mortgage-Backed Securities Purchase Program," *International Journal of Central Banking*, 8, 1–42.

TROLLE, A. B. AND E. S. SCHWARTZ (2009): "A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives," *Review of Financial Studies*, 22, 2007–2057.

VICKERY, J. AND J. WRIGHT (2013): "TBA Trading and Liquidity in the Agency MBS Market," *Federal Reserve Bank of New York Economic Policy Review*, 19, 1–18.

WOODFORD, M. (2012): "Methods of Policy Accommodation at the Interest-Rate Lower Bound," Presented at the Federal Reserve Bank of Kansas City's Jackson Hole Symposium.

Figure 1: Share of total MBS value by coupon. Each shaded grey area represents the share (left axis) of total balance in the 30-year fixed rate mortgage universe accounted for by MBS with a given coupon (heat map in the right panel). The red line is the current coupon TBA (right axis).

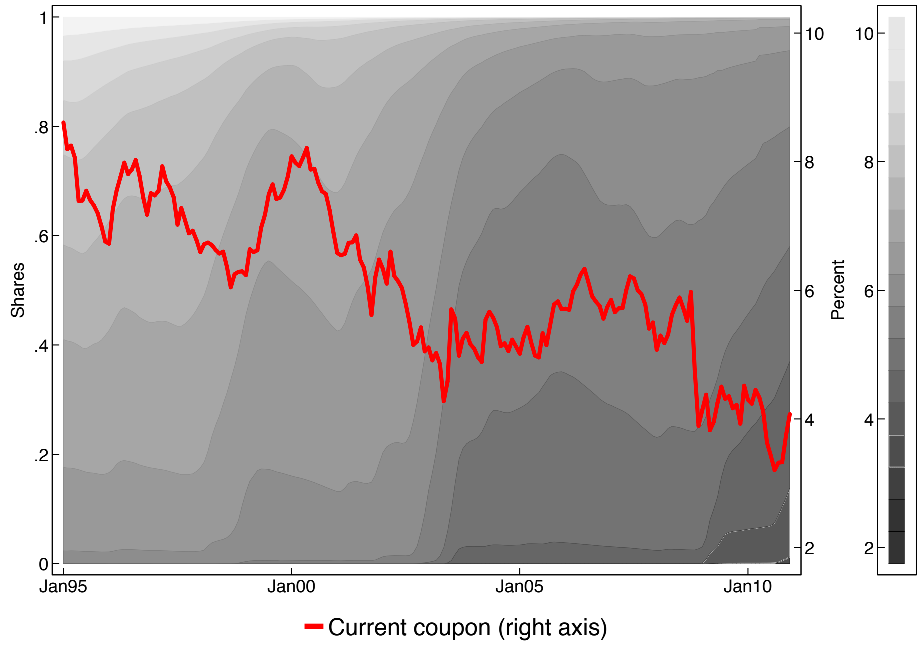
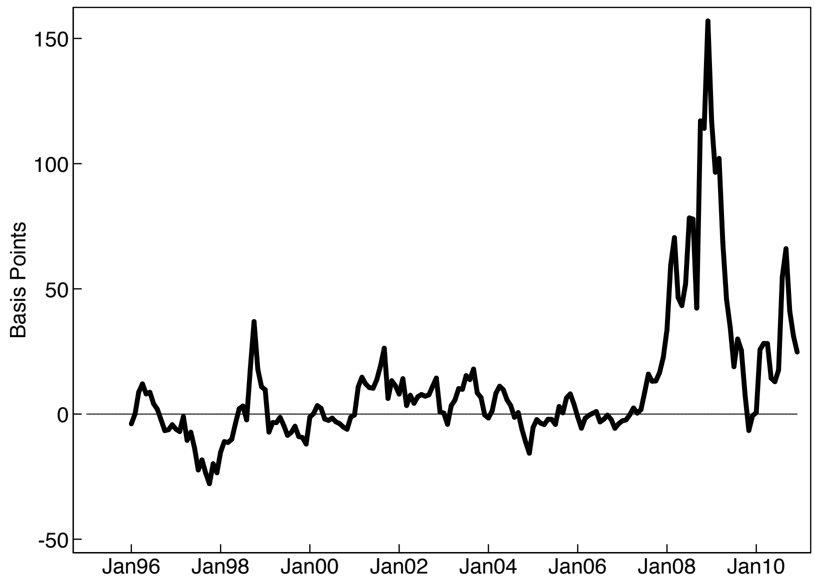


Figure 2: Time-series and cross-sectional variation in OAS of MBS in TBA market. The top panel shows the time-series variation in the option-adjusted spread (to swaps) on a value-weighted index based on TBA quotes from six dealers. The bottom panel displays a scatterplot and a local smoother of the cross-sectional variation in OAS for MBS coupons with remaining principal balance (in 2009 dollars) of $100 million or more. Further details reported in Section 2.

(a) Time Series



(b) Cross Section

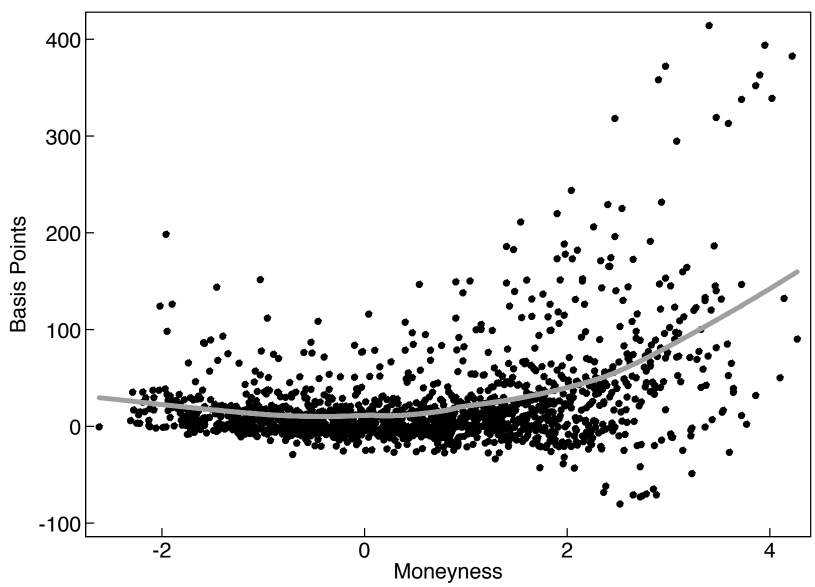


Figure 3: Cross-sectional variation in hedged returns. The figure shows one-year cumulative hedged returns (over months t to $t+12$) on MBS in the Barclays index against securities' moneyness in month t. Only includes securities for which 12 consecutive months of returns are available; other data cleaning as for TBA analysis (see Appendix B).

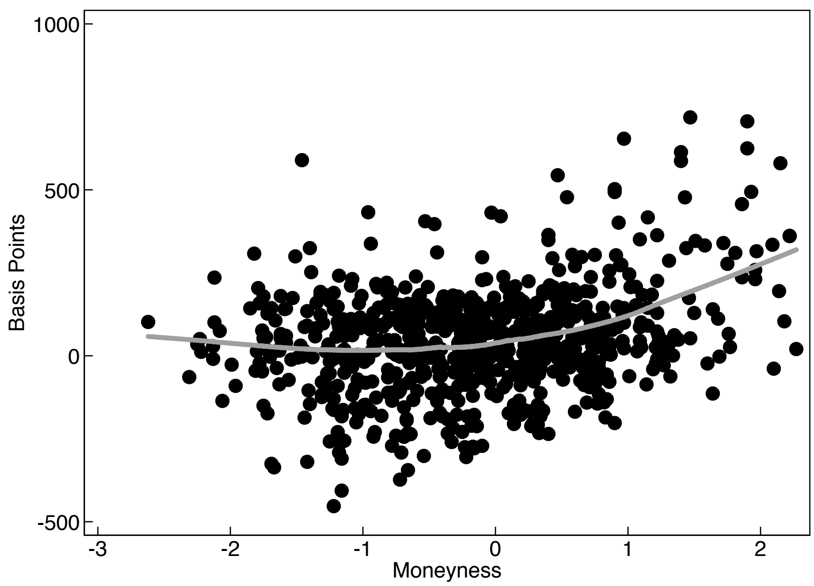


Figure 4: MBS cash flows for different prepayment speeds. The colored areas represent cash flows for a hypothetical MBS with original principal of \$100, loan rate of 4% and coupon of 4.5% in a slow (top panel) and a fast (bottom panel) prepayment scenario.

(a) Slow prepayment (CPR = 12%)

(b) Fast prepayment (CPR = 24%)

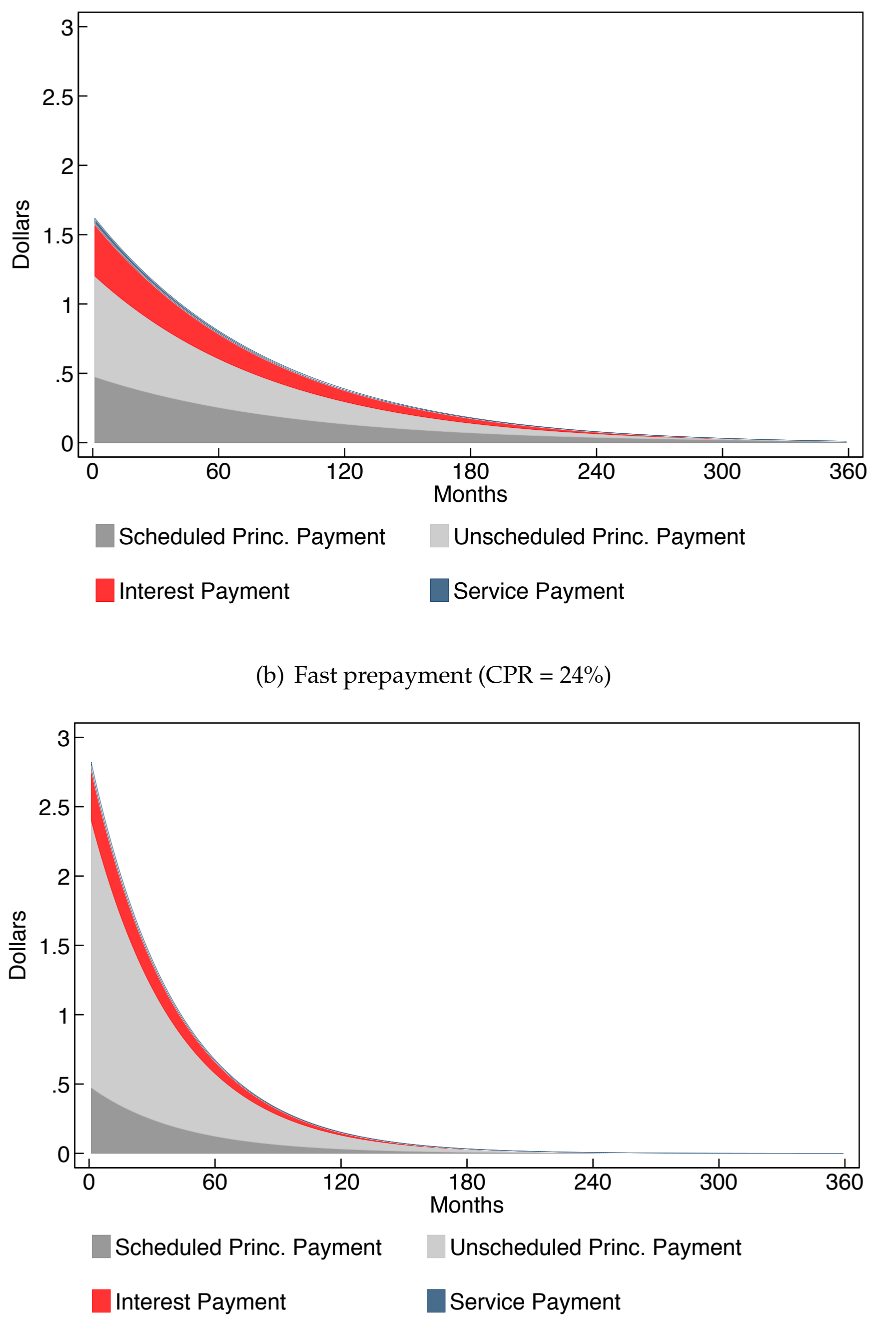


Figure 5: Graphical explanation of the model identification. The figure shows OAS on the IO, PO and pass-through as a function of the multiple (Λ) on the physical prepayment speed. The OAS on the IO (PO) declines (increases) in Λ. The OAS on the pass-through in this example also declines in Λ because the pass-through is a premium security ($P^{IO} + P^{PO} > \$100$). The three OAS differ at the physical speed (Λ) but are equalized and equal to $OAS^{\mathbb{Q}}$ at the risk-neutral speed.

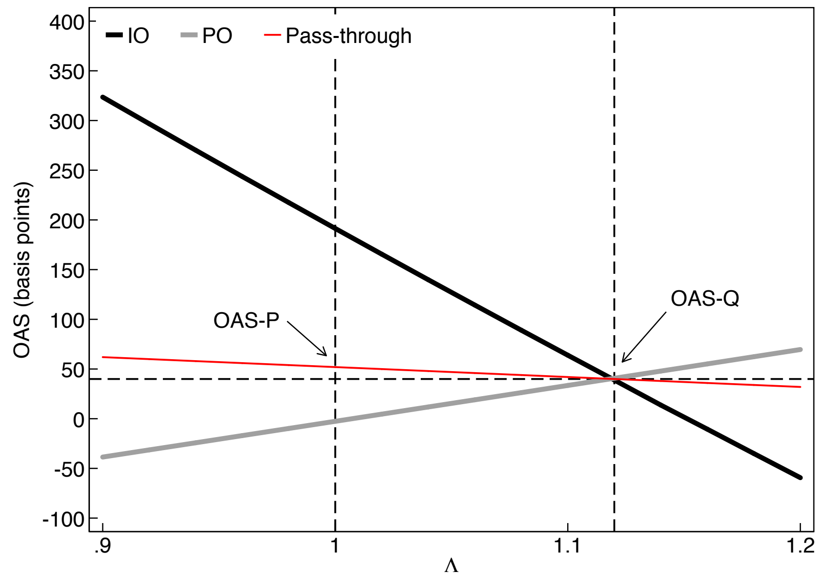


Figure 6: S curve for different pool compositions. The figure shows the prepayment function in equation (4.1) with parameters b_1, b_2 and b_3 equal to their averages across securities in our sample for different levels of the fraction χ of borrowers with high interest rate sensitivity. The vertical axis is the annualized "conditional prepayment rate" (CPR) and the horizontal axis is the incentive to prepay.

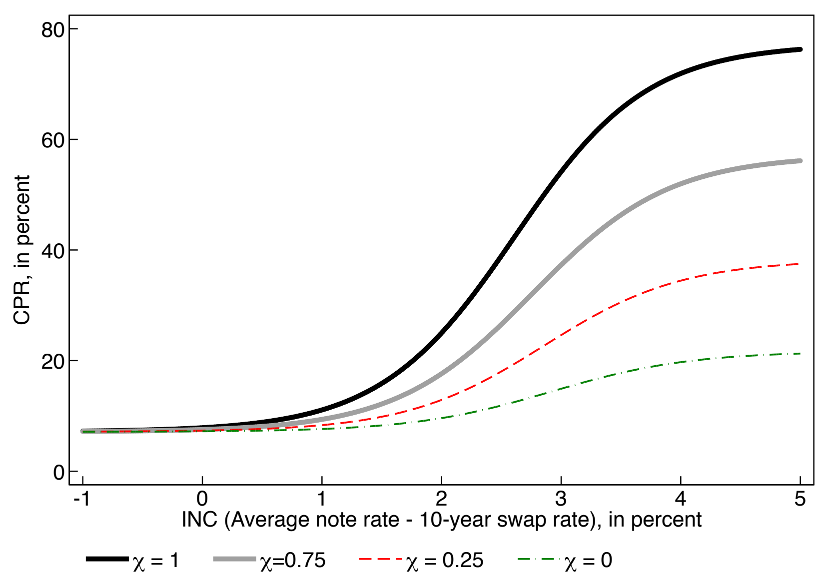


Figure 7: Cross-sectional variation in spreads on pass-throughs. The panels show scatterplots and local smoothers of the cross-sectional variation in OAS, $OAS^{\mathbb{Q}}$, and the difference between the physical and market-implied prepayment speed using IO/PO prices and our pricing model. Additional detail is available in Section 5.

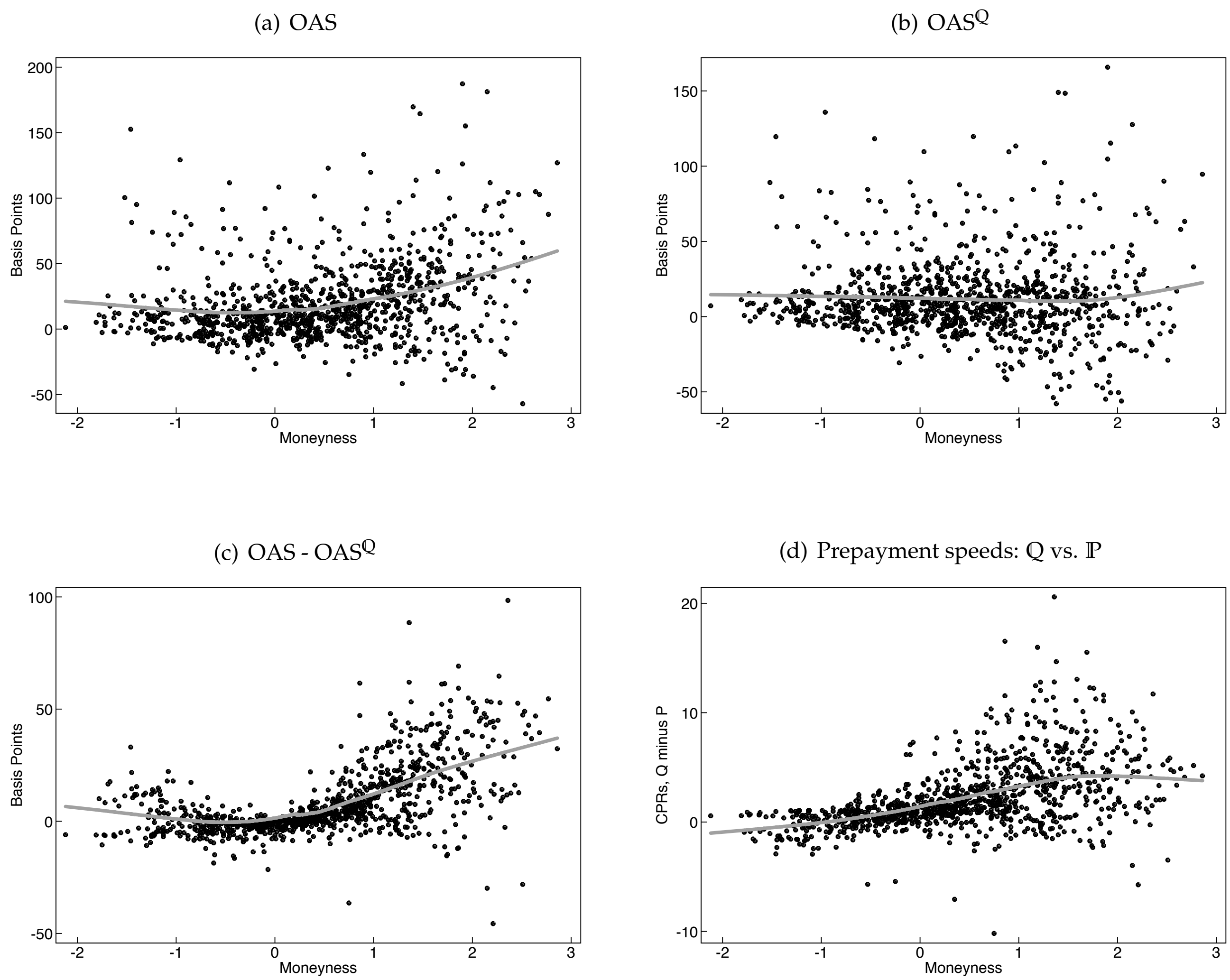


Figure 8: Time series of spreads on the pass-through index. This figure shows time-series variation in OAS and $OAS^{\mathbb{Q}}$ on a value-weighted index based on IO/PO prices and our pricing model. Additional detail is available in Section 5.

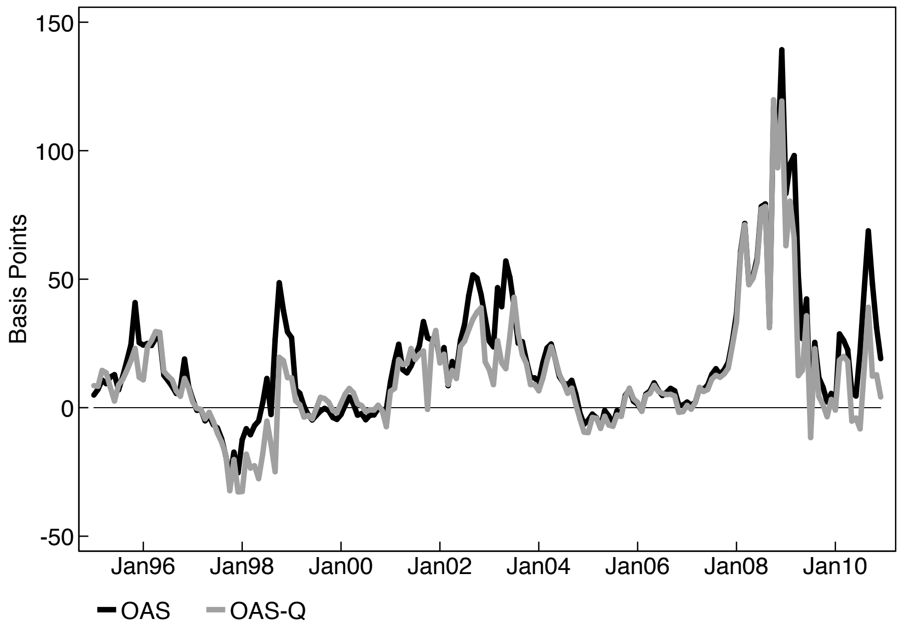


Figure 9: OAS around initial LSAP announcement. This figure shows OAS variation for the four main coupons from August 2008 to May 2009, based on median TBA quotes from dealers.

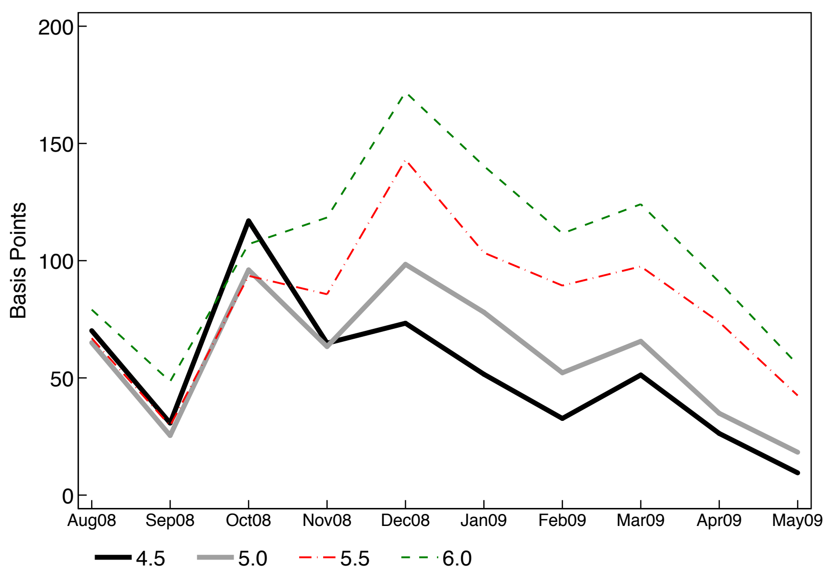
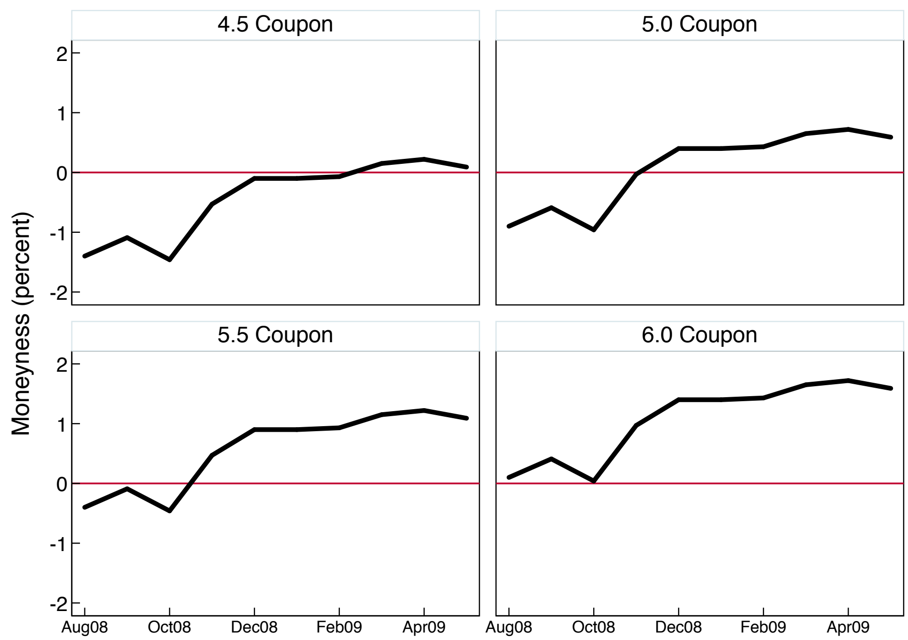


Figure 10: OAS decomposition around initial LSAP announcement This figure shows the MBS moneyness by coupon (upper panels) and movements in OAS and $OAS^{\mathbb{Q}}$ (bottom panels) based on IO/PO quotes and our prepayment model.

(a) Moneyness across coupons



(b) Spreads across coupons

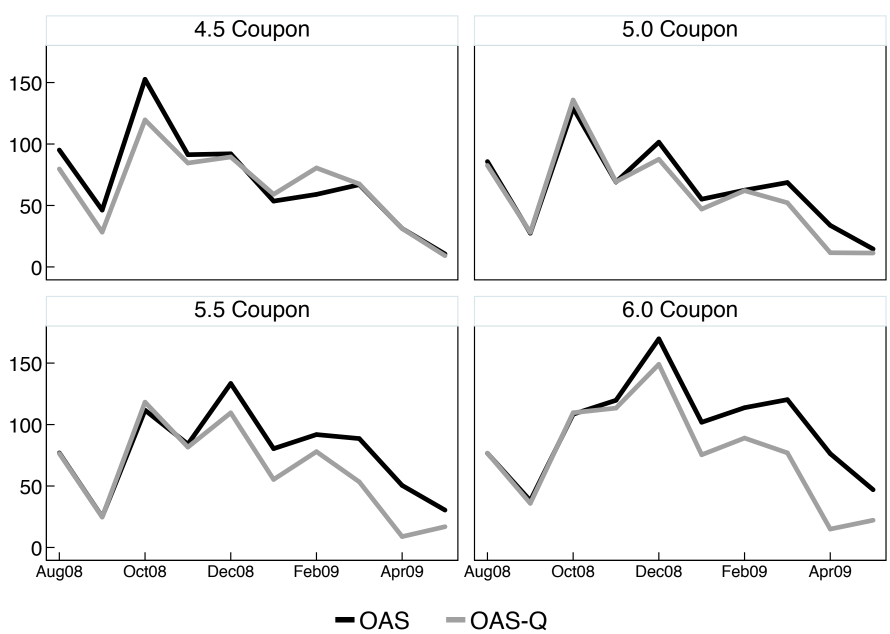


Table 1: Time-series regressions of OAS on TBA index. Coefficient estimates from OLS regression of the TBA index OAS on the Aaa-Treasury spread, the Baa-Aaa spread, the 10-year to 3-month slope of the Treasury curve and the 2-year into 10-year swaption implied volatility. All variables are measured end-of-month. Regressors are standardized to have zero mean and unit standard deviation. Newey-West standard errors (6 lags) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)		(2)	
Aaa - Treas	3.2**	[1.5]	2.3*	[1.2]
Baa - Aaa	19.6***	[3.4]	9.7**	[3.8]
Treas Slope	-0.5	[2.4]	2.0	[1.3]
Swaption Vol.	2.7	[4.2]	-0.8	[2.4]
Const	8.8***	[1.8]	3.6***	[1.3]
Adj. R2	0.76		0.31	
Obs.	180		139	
Dates	199601.201012		199601.200707	

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

	(1)		(2)		(3)		(4)	
< -1.75	18.7***	[4.5]	18.0***	[2.5]	14.9***	[2.0]	14.1***	[2.5]
$[-1.75, -1.25)$	8.9***	[1.9]	10.4***	[1.7]	7.7***	[1.2]	8.3***	[1.5]
$[-1.25, -.75)$	4.0***	[1.0]	5.3***	[1.2]	4.3***	[0.9]	4.8***	[0.9]
$[-.75, -.25)$	-0.0	[0.4]	-0.1	[0.4]	0.8*	[0.4]	0.8*	[0.4]
$[.25, .75)$	1.5***	[0.5]	1.5***	[0.6]	-0.1	[0.4]	-0.1	[0.4]
$[.75, 1.25)$	6.4***	[1.2]	6.4***	[1.3]	2.6***	[0.9]	2.5***	[1.0]
$[1.25, 1.75)$	15.6***	[2.2]	15.0***	[2.2]	7.2***	[1.8]	7.2***	[1.9]
$[1.75, 2.25)$	27.2***	[3.7]	24.6***	[3.6]	11.9***	[3.0]	11.4***	[3.2]
$[2.25, 2.75)$	41.3***	[6.4]	33.9***	[5.8]	26.0***	[5.3]	23.2***	[5.5]
≥ 2.75	93.4***	[13.3]	85.4***	[11.6]	50.5***	[7.2]	47.1***	[7.6]
Const	7.5***	[1.5]	8.6***	[1.5]	0.5	[0.8]	0.8	[0.8]
Month FEs?	No		Yes		No		Yes	
Adj. R2	0.22		0.61		0.24		0.38	
Obs.	1532		1532		1151		1151	
Dates	199601.201012		199601.201012		199601.200707		199601.200707	

Table 3: Regressions of one-year returns on OAS. Coefficient estimates from OLS regression of the cumulative (hedged or unhedged) MBS return (in percent) on different coupons over months t to $t + 12$ on OAS (averaged across dealers) in month t, either including or excluding time fixed effects. Only includes observations where 12 consecutive monthly returns are available. Newey-West standard errors (12 lags) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) Hedged		(2) Hedged		(3) Unhedged		(4) Unhedged	
Coupon OAS	3.3^{***}	[0.4]	3.8^{***}	[0.6]	1.2^{*}	[0.6]	3.1^{***}	[1.0]
Const	0.1	[0.1]	1.5^{***}	[0.1]	6.6^{***}	[0.4]	4.8^{***}	[0.6]
Month FEs?	No		Yes		No		Yes	
Adj. R2	0.34		0.75		0.01		0.87	
Obs.	734		734		734		734	
Dates	199601.201012		199601.201012		199601.201012		199601.201012	

Table 4: Cross section of OAS$^{\mathbb{Q}}$ on pass-throughs. Coefficient estimates from OLS regression of the $OAS^{\mathbb{Q}}$ from our model on different moneyness level bins either including or excluding time fixed effects. Robust standard errors (clustered at the month level) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)		(2)		(3)		(4)	
< -1.25	1.5	[3.8]	-1.3	[1.5]	-1.8	[1.5]	-1.5	[1.7]
$[-1.25, -.75)$	-0.2	[2.1]	-2.7**	[1.2]	-2.9**	[1.3]	-3.3**	[1.3]
$[-.75, -.25)$	-1.4	[1.3]	-2.3**	[0.9]	-1.2	[1.0]	-2.5***	[0.9]
$[.25, .75)$	-2.3***	[0.7]	-2.5***	[0.8]	-1.0	[0.8]	-1.5**	[0.8]
$[.75, 1.25)$	-5.2***	[1.6]	-4.0***	[1.5]	-3.6**	[1.4]	-4.1***	[1.5]
$[1.25, 1.75)$	-3.3	[2.5]	-1.4	[2.4]	-4.6*	[2.3]	-4.9*	[2.6]
$[1.75, 2.25)$	-3.1	[4.6]	2.0	[3.8]	-9.7**	[4.4]	-4.4	[4.1]
≥ 2.25	6.1	[6.9]	16.5***	[6.3]	9.1	[8.6]	15.5*	[8.1]
Const	13.8***	[1.6]	13.1***	[0.9]	7.9***	[1.1]	7.9***	[0.8]
Month FEs?	No		Yes		No		Yes	
Adj.	-0.00		0.73		0.02		0.52	
Obs.	1005		1005		796		796	
Dates	199501.201012		199501.201012		199501.200707		199501.200707	

Table 5: Cross section of OAS$^{\mathbb{P}}$ - OAS$^{\mathbb{Q}}$ on pass-throughs. Coefficient estimates from OLS regression of the prepayment-risk component in the OAS from our model on different moneyness level bins either including or excluding time fixed effects. Robust standard errors (clustered at the month level) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)		(2)		(3)		(4)	
< -1.25	5.4***	[1.9]	6.8***	[2.2]	2.3	[1.7]	3.3*	[2.0]
$[-1.25, -.75)$	2.9***	[0.9]	5.2***	[1.2]	0.9	[0.8]	3.2***	[1.1]
$[-.75, -.25)$	-1.0***	[0.4]	0.7	[0.7]	-1.8***	[0.3]	-0.0	[0.7]
$[.25, .75)$	4.9***	[0.4]	4.1***	[0.5]	5.1***	[0.4]	4.5***	[0.5]
$[.75, 1.25)$	13.3***	[1.0]	12.4***	[1.1]	13.5***	[1.1]	12.8***	[1.3]
$[1.25, 1.75)$	22.1***	[1.5]	20.7***	[1.7]	22.3***	[1.7]	21.2***	[1.9]
$[1.75, 2.25)$	26.6***	[2.7]	23.1***	[2.7]	26.5***	[3.5]	23.3***	[3.5]
≥ 2.25	34.2***	[5.0]	27.5***	[4.3]	30.2***	[6.1]	25.6***	[5.1]
Const	-1.4***	[0.2]	-1.0	[0.7]	-1.5***	[0.3]	-1.2*	[0.7]
Month FEs?	No		Yes		No		Yes	
Adj. R2	0.42		0.59		0.42		0.58	
Obs.	1005		1005		796		796	
Dates	199501.201012		199501.201012		199501.200707		199501.200707	

Table 6: Time-series regressions on pass-throughs (from stripped MBS). Coefficient estimates from OLS regression of spreads reported in the top row of each panel on the Baa-Aaa spread, MBS issuance to equity of brokers and dealers, average moneyness in the MBS universe squared and spreads on agency debt to swaps. All regressors are standardized to have zero mean and unit standard deviation. Newey-West standard errors (6 lags) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) OAS		(2) OAS^Q		(3) OAS - OAS^Q		(4) OAS		(5) OAS^Q		(6) OAS - OAS^Q	
Baa - Aaa	14.2***	[2.7]	11.7***	[3.0]	2.5***	[0.7]	11.9***	[4.0]	10.1*	[5.2]	1.8	[2.5]
MBS Issuance / BD Equity	6.7***	[1.1]	5.1***	[1.2]	1.6**	[0.8]	6.1***	[1.3]	3.8*	[2.1]	2.2	[1.4]
Mean Moneyness^2	2.2	[1.9]	-4.6**	[2.0]	6.8***	[1.0]	4.5***	[1.6]	-1.3	[2.0]	5.8***	[1.4]
Agency Debt Spreads	6.9***	[2.1]	8.0***	[2.3]	-1.0	[1.1]	8.1***	[2.0]	8.9***	[2.3]	-0.8	[1.3]
Const	17.2***	[1.7]	11.8***	[1.9]	5.4***	[0.8]	15.7***	[1.4]	10.8***	[1.7]	4.8***	[1.2]
Adj. R2	0.71		0.63		0.57		0.68		0.50		0.46	
Obs.	192		192		192		151		151		151	
Dates	199501.201012		199501.201012		199501.201012		199501.200707		199501.200707		199501.200707	

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 θ_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 inclusive of interest and principal 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, s, or by the guaranteeing agency, g: $s + g = c - w$. 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:

$$\tilde{\theta}_t = (1 + w)\,\tilde{\theta}_{t-1} - k. \tag{A.1}$$

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 θ_t including early prepayment generalizes equation (A.1) to take into account prepayments. After accounting for loan amortization and unscheduled principal payments, the factor evolves according to:

$$\theta_t = (1 - \mathrm{SMM}_{t-1})(1 + w)\theta_{t-1} - k\,\hat{\theta}_t, \tag{A.2}$$

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, $\hat{\theta}_t = \prod_{s=0}^{t-1}(1 - \mathrm{SMM}_s)$. The prepayment speed is often reported in annualized terms, known as the "conditional prepayment rate" or $\mathrm{CPR}_t = 1 - (1 - \mathrm{SMM}_t)^{12}$. Given prepayment rates, cash flows passed through to investors per unit of principal are:

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

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 $c\,\theta_{t-1}$.

MBS cash flows shown in equation (A.3) are obtained from the path of θ_t by noting that the scheduled payment $(\theta_{t-1} - \theta_t)$ is equal to the total of the scheduled component $(k\hat{\theta}_t - w\theta_{t-1})$ and the unscheduled one $(SMM_{t-1}\,(1 + w)\,\theta_{t-1})$. Finally, borrowers pay $w\theta_{t-1}$ interest to servicers, which pass through the payments to investors after keeping a servicing stream equal to $(c - w)\theta_{t-1}$ net of a fraction that is used to pay the guarantee fee to the agency.

B Data

B.1 TBA sample and data-quality filters

The sample spans 1996 to 2010, reflecting limited data availability on TBAs, which we use to characterize the facts in this section, prior to 1996, and a limited liquidity in IO/PO strips, which we use later in the paper to decompose the OAS, post 2010. 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.

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 don't 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 Expected return derivations

In this Appendix, we derive (2.3) and show that the expected hedged return under the interest-rate-risk-neutral measure $\mathbb{Q}^*$ equals the expected hedged return under the physical measure $\mathbb{P}$.

We begin by showing that the OAS can be approximated as a weighted average of future excess returns. Approximating around 0 excess return

$$\left.\frac{\partial \prod_{j=1}^{k}\left(1+r_{t+j}+\mathbb{E}_{t+j-1}^{*}\left[rx_{t+j}\right]\right)^{-1}}{\partial \mathbb{E}_{t+j-1}^{*}\left[rx_{t+j}\right]}\right|_{\mathbb{E}_{t+j-1}^{*}\left[rx_{t+j}\right]=0} = -\frac{1}{1+r_{t+j}}\prod_{j=1}^{k}\frac{1}{\left(1+r_{t+j}\right)}$$

$$\left.\frac{\partial \prod_{j=1}^{k}\left(1+r_{t+j}+OAS_t\right)^{-1}}{\partial OAS_t}\right|_{OAS_t=0} = -\sum_{l=1}^{k}\frac{1}{1+r_{t+l}}\prod_{j=1}^{k}\frac{1}{\left(1+r_{t+j}\right)},$$

(2.2) becomes

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

Similarly, (2.1) becomes

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

Comparing these two expressions, we obtain

$$OAS_t\mathbb{E}_t^{*}\left[\sum_{k=1}^{T-t}\frac{X_k\left(r_{t+k}\right)}{\prod_{j=1}^{k}\left(1+r_{t+j}\right)}\sum_{l=1}^{k}\frac{1}{1+r_{t+l}}\right] = \mathbb{E}_t^{*}\left[\sum_{k=1}^{T-t}\frac{X_k\left(r_{t+k}\right)}{\prod_{j=1}^{k}\left(1+r_{t+j}\right)}\sum_{l=1}^{k}\frac{\mathbb{E}_{t+l-1}^{*}\left[rx_{t+l}\right]}{1+r_{t+l}}\right].$$

Thus, the OAS is a weighted average of expected excess one period returns over the lifetime of the security. Rearranging, we obtain (2.3).

Using the fund separation theorem (see e.g. Merton, 1972), we can represent the return on the

MBS as

$$
\begin{aligned}
rx_{t+1} &= \sum_{j=1}^{N_r} \frac{\mathbb{C}_t\left(rx_{t+1}, z_{t+1}^j\right)}{\mathbb{V}_t\left(z_{t+1}^j\right)} z_{t+1}^j + \sum_{j=1}^{N_p} \frac{\mathbb{C}_t\left(rx_{t+1}, p_{t+1}^j\right)}{\mathbb{V}_t\left(p_{t+1}^j\right)} p_{t+1}^j + \epsilon_{t+1} \\
&\equiv \sum_{j=1}^{N_r} \beta_{rj,t}^i z_{t+1}^j + \sum_{j=1}^{N_p} \beta_{pj,t}^i p_{t+1}^j + \epsilon_{t+1},
\end{aligned}
$$

where $\left\{z_{t+1}^j\right\}_{j=1}^{N_r}$ are the one period excess returns on a set of portfolios that spans interest rate uncertainty, and $\left\{p_{t+1}^j\right\}_{j=1}^{N_p}$ are the one-period excess returns on a set of portfolios that spans the prepayment risk that is orthogonal to interest rates. The return on the hedged portfolio is then given by

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

Then the expected excess return under $\mathbb{Q}$ is given by

$$
\mathbb{E}_t^*\left[rx_{t+1} - \sum_{j=1}^{N_r} \beta_{rj,t}^i z_{t+1}^j\right] = \mathbb{E}_t^*\left[\sum_{j=1}^{N_p} \beta_{pj,t}^i p_{t+1}^j + \epsilon_{t+1}\right] = \mathbb{E}_t\left[\mathcal{M}_{rt}\left(\sum_{j=1}^{N_p} \beta_{pj,t}^i p_{t+1}^j + \epsilon_{t+1}\right)\right],
$$

where $\mathcal{M}_{r,t+1}$ is the change of measure between $\mathbb{P}$ and $\mathbb{Q}^*$. Since $\left\{z_{t+1}^j\right\}_{j=1}^{N_r}$ span the interest rate uncertainty, $\mathcal{M}_{r,t+1}$ can be represented as a portfolio of $\left\{z_{t+1}^j\right\}_{j=1}^{N_r}$

$$
\mathcal{M}_{r,t+1} = m_{0t} + \sum_{j=1}^{N_r} m_{jt} z_{t+1}^j,
$$

with $1 = m_{0t} + \sum_{j=1}^{N_r} m_{jt} \mathbb{E}_t\left[z_{t+1}^j\right]$. Thus

$$
\begin{aligned}
\mathbb{E}_t^*\left[rx_{t+1} - \sum_{j=1}^{N_r} \beta_{rj,t}^i z_{t+1}^j\right] &= \mathbb{E}_t\left[\left(m_{0t} + \sum_{j=1}^{N_r} m_{jt} z_{t+1}^j\right)\left(\sum_{j=1}^{N_p} \beta_{pj,t}^i p_{t+1}^j + \epsilon_{t+1}\right)\right] \\
&= \mathbb{E}_t\left[m_{0t} + \sum_{j=1}^{N_r} m_{jt} z_{t+1}^j\right] \mathbb{E}_t\left[\sum_{j=1}^{N_p} \beta_{pj,t}^i p_{t+1}^j + \epsilon_{t+1}\right] \\
&= \mathbb{E}_t\left[\sum_{j=1}^{N_p} \beta_{pj,t}^i p_{t+1}^j + \epsilon_{t+1}\right],
\end{aligned}
$$

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 $\mathcal{M}_{r,t+1}$ is a martingale.

D Additional details on model

In this Appendix, we derive the formulas for OAS on the pass-through security used in Section 3, as well as on the IO and PO strips. Recall that, in each instant of time dt, the investor in the pass-through security on pool j receives $c_j - d\theta_{jt}$ per dollar of face value. Thus, the price of one unit of the pass-through satisfies

$$P^j_{PT,t} = \mathbb{E}_t \left[\int_0^{+\infty} \frac{M_{t+s}}{M_t} \left(c_j + s_{js} \right) ds \right] = \mathbb{E}^{\mathbb{Q}}_t \left[\int_0^{+\infty} e^{-\int_0^s \left(r_u + s_{ju} \right) du} \left(c_j + s_{js} \right) ds \right],$$

where $\mathbb{Q}$ is the risk-neutral measure associated with the pricing kernel M. Assume that the short rate r_t evolves according to

$$dr_t = \mu_{rt} dt + \sigma_r dZ_{rt},$$

where Z_{rt} is a standard Brownian motion, independent of the shocks to the prepayment function parameters, $Z_{\gamma t}$. Then, applying the Feynman-Kac theorem, we can represent the price of the pass-through security as the solution to

$$\begin{aligned} r_t P^j_{PT,t} = \left(c_j + s_{jt} \right) - s_{jt} P^j_{PT,t} - \mu_t \pi_{lt} \alpha_j P^j_{PT,t} + \frac{\partial P^j_{PT,t}}{\partial r_t} \left(\mu_{rt} + \sigma_r \pi_{rt} \right) + \frac{1}{2} \frac{\partial^2 P^j_{PT,t}}{\partial r_t^2} \sigma_r^2 \\ + \frac{\partial P^j_{PT,t}}{\partial \gamma_t'} \left(\mu_\gamma + \sigma_\gamma \pi_{\gamma t} \right) + \frac{1}{2} \mathrm{tr} \left(\frac{\partial^2 P^j_{PT,t}}{\partial \gamma_t \partial \gamma_t'} \sigma_\gamma \sigma_\gamma' \right), \end{aligned} \tag{D.1}$$

where π_{rt} is the price of risk associated with innovations to the short rate Z_{rt}.

The option-adjusted spread (OAS) is computed using the risk-neutral evolution of the interest rates, but ignoring variation of non-interest rate sources of risk. More formally, the OAS on the pass-through solves

$$\begin{aligned} \left(r_t + OAS_{jt} \right) P^j_{PT,t} = \left(c_j + s_{jt} \right) - s_{jt} P^j_{PT,t} - \mu_t \alpha_j P^j_{PT,t} + \frac{\partial P^j_{PT,t}}{\partial r_t} \left(\mu_{rt} + \sigma_r \pi_{rt} \right) + \frac{\partial^2 P^j_{PT,t}}{\partial r_t^2} \frac{\sigma_r^2}{2} \\ + \frac{\partial P^j_{PT,t}}{\partial \gamma_t'} \mu_\gamma + \frac{1}{2} \mathrm{tr} \left(\frac{\partial^2 P^j_{PT,t}}{\partial \gamma_t \partial \gamma_t'} \sigma_\gamma \sigma_\gamma' \right). \end{aligned} \tag{D.2}$$

Comparing (D.1) and (D.2), we see that the OAS on the pass-through is given by

$$OAS_{jt} = -\pi_{\gamma t}' \sigma_\gamma \frac{1}{P^j_{PT,t}} \frac{\partial P^j_{PT,t}}{\partial \gamma_t} + \alpha_j \mu_t \left(\pi_{lt} - 1 \right).$$

Thus, the OAS is the risk premium paid to the MBS investors for holding non-interest rate risk.

We compute $OAS^{\mathbb{Q}}$ taking into account compensation for prepayment risk, so that the $OAS^{\mathbb{Q}}$

solves

$$\left(r_t + OAS^{\mathbb{Q}}_{jt}\right) P^j_{PT,t} = \left(c_j + s_{jt}\right) - s_{jt} P^j_{PT,t} - \mu_t \alpha_j P^j_{PT,t} + \frac{\partial P^j_{PT,t}}{\partial r_t}\left(\mu_{rt} + \sigma_r \pi_{rt}\right) + \frac{\partial^2 P^j_{PT,t}}{\partial r_t^2}\frac{\sigma_r^2}{2} + \frac{\partial P^j_{PT,t}}{\partial \gamma_t'}\left(\mu_\gamma + \sigma_\gamma \pi_{\gamma t}\right) + \frac{1}{2}\mathrm{tr}\left(\frac{\partial^2 P^j_{PT,t}}{\partial \gamma_t \partial \gamma_t'}\sigma_\gamma \sigma_\gamma'\right). \quad \text{(D.3)}$$

Comparing (D.1) and (D.3), we see that the $OAS^{\mathbb{Q}}$ on the pass-through is given by

$$OAS^{\mathbb{Q}}_{jt} = \alpha_j \mu_t \left(\pi_{lt} - 1\right).$$

We can compute the OAS and $OAS^{\mathbb{Q}}$ on stripped securities in a similar fashion. In each instant of time dt, an investor holding the IO strip on pool j receives the coupon payments c_j while an investor holding the PO strip receives the principal payments $-d\theta_{jt}$. Then, similarly to (D.1), the price $P^j_{IO,t}$ of the IO strip is the solution to

$$r_t P^j_{IO,t} = c_j - s_{jt} P^j_{IO,t} - \mu_t \pi_{lt} \alpha_j P^j_{IO,t} + \frac{\partial P^j_{IO,t}}{\partial r_t}\left(\mu_{rt} + \sigma_r \pi_{rt}\right) + \frac{1}{2}\frac{\partial^2 P^j_{IO,t}}{\partial r_t^2}\sigma_r^2 + \frac{\partial P^j_{IO,t}}{\partial \gamma_t'}\left(\mu_\gamma + \sigma_\gamma \pi_{\gamma t}\right) + \frac{1}{2}\mathrm{tr}\left(\frac{\partial^2 P^j_{IO,t}}{\partial \gamma_t \partial \gamma_t'}\sigma_\gamma \sigma_\gamma'\right), \quad \text{(D.4)}$$

and the price $P^j_{PO,t}$ of the PO strip is the solution to

$$r_t P^j_{PO,t} = s_{jt} - s_{jt} P^j_{PO,t} - \mu_t \pi_{lt} \alpha_j P^j_{PO,t} + \frac{\partial P^j_{PO,t}}{\partial r_t}\left(\mu_{rt} + \sigma_r \pi_{rt}\right) + \frac{1}{2}\frac{\partial^2 P^j_{PO,t}}{\partial r_t^2}\sigma_r^2 + \frac{\partial P^j_{PO,t}}{\partial \gamma_t'}\left(\mu_\gamma + \sigma_\gamma \pi_{\gamma t}\right) + \frac{1}{2}\mathrm{tr}\left(\frac{\partial^2 P^j_{PO,t}}{\partial \gamma_t \partial \gamma_t'}\sigma_\gamma \sigma_\gamma'\right). \quad \text{(D.5)}$$

The OAS and $OAS^{\mathbb{Q}}$ on the stripped securities are defined analogously to (D.2)-(D.3). Thus, OAS and $OAS^{\mathbb{Q}}$ on the IO strip are given, respectively, by

$$OAS_{jt,IO} = -\pi_{\gamma t}' \sigma_\gamma \frac{1}{P^j_{IO,t}} \frac{\partial P^j_{IO,t}}{\partial \gamma_t} + \alpha_j \mu_t \left(\pi_{lt} - 1\right)$$
$$OAS^{\mathbb{Q}}_{jt,IO} = \alpha_j \mu_t \left(\pi_{lt} - 1\right).$$

Similarly, the OAS and $OAS^{\mathbb{Q}}$ on the PO strip are given, respectively, by

$$\begin{aligned} OAS_{jt,PO} &= -\pi'_{\gamma t}\sigma_{\gamma}\frac{1}{P^{j}_{PO,t}}\frac{\partial P^{j}_{PO,t}}{\partial \gamma_t} + \alpha_j \mu_t \left(\pi_{It} - 1\right) \\ OAS^{\mathbb{Q}}_{jt,PO} &= \alpha_j \mu_t \left(\pi_{It} - 1\right). \end{aligned}$$

Thus, when the stripped MBS have equal exposure to liquidity risk, the prepayment risk-neutral OAS, $OAS^{\mathbb{Q}}$, is equal for an IO and PO pair, as well as the corresponding pass-through.

E Pricing model details

E.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

$$df(t,T) = \mu_f(t,T)\,dt + \sum_{i=1}^{3}\sigma_{f,i}(t,T)\,dW^{\mathbb{Q}}_{i,t},$$

where $W^{\mathbb{Q}}_{it}$ are independent standard Weiner processes under the risk-neutral measure $\mathbb{Q}$, 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 $\sigma_{f,i}(t,T)$ is

$$\sigma_{f,i}(t,T) = \left(\alpha_{0,i} + \alpha_{1,i}(T-t)\right)e^{-\gamma_i(T-t)}. \tag{E.1}$$

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 (E.1), the time t price of a zero-coupon bond maturing at time T, $P(t,T)$, is given by

$$P(t,T) \equiv \exp\left\{-\int_t^T f(t,u)\,du\right\} = \frac{P(0,T)}{P(0,t)}\exp\left\{\sum_{i=1}^{3}B_{x_i}(T-t)\,x_{it} + \sum_{i=1}^{3}\sum_{j=1}^{6}B_{\phi_{ji}}(T-t)\,\phi_{ji,t}\right\},$$

where the state variables $\left\{x_{it}, \phi_{ji,t}\right\}$ follow

$$\begin{aligned}
dx_{it} &= -\gamma_i x_{it} dt + dW_{it}^{\mathbb{Q}} \\
d\phi_{1i,t} &= \left(x_{it} - \gamma_i \phi_{1i,t}\right) dt \\
d\phi_{2i,t} &= \left(1 - \gamma_i \phi_{2i,t}\right) dt \\
d\phi_{3i,t} &= \left(1 - 2\gamma_i \phi_{3i,t}\right) dt \\
d\phi_{4i,t} &= \left(\phi_{2i,t} - \gamma_i \phi_{4i,t}\right) dt \\
d\phi_{5i,t} &= \left(\phi_{3i,t} - 2\gamma_i \phi_{5i,t}\right) dt \\
d\phi_{6i,t} &= \left(2\phi_{5i,t} - 2\gamma_i \phi_{6i,t}\right) dt.
\end{aligned}$$

The coefficients $\left\{B_{x_i}, B_{\phi_{ji}}\right\}$ are functions of the parameters of the volatility function and the time to maturity $\tau = T - t$, and are given by

$$\begin{aligned}
B_{x_i}(\tau) &= \frac{\alpha_{1i}}{\gamma_i}\left(\left(\frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}}\right)\left(e^{-\gamma_i\tau} - 1\right) + \tau e^{-\gamma_i\tau}\right) \\
B_{\phi_{1i}}(\tau) &= \frac{\alpha_{1i}}{\gamma_i}\left(e^{-\gamma_i\tau} - 1\right) \\
B_{\phi_{2i}}(\tau) &= \left(\frac{\alpha_{1i}}{\gamma_i}\right)^2\left(\frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}}\right)\left(\left(\frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}}\right)\left(e^{-\gamma_i\tau} - 1\right) + \tau e^{-\gamma_i\tau}\right) \\
B_{\phi_{3i}}(\tau) &= -\frac{\alpha_{1i}}{\gamma_i^2}\left(\left(\frac{\alpha_{1i}}{2\gamma_i^2} + \frac{\alpha_{0i}}{\gamma_i} + \frac{\alpha_{0i}^2}{2\alpha_{1i}}\right)\left(e^{-2\gamma_i\tau} - 1\right) + \left(\frac{\alpha_{1i}}{\gamma_i} + \alpha_{0i}\right)\tau e^{-2\gamma_i\tau} + \frac{\alpha_{1i}}{2}\tau^2 e^{-2\gamma_i\tau}\right) \\
B_{\phi_{4i}}(\tau) &= \left(\frac{\alpha_{1i}}{\gamma_i}\right)^2\left(\frac{1}{\gamma_i} + \frac{\alpha_{0i}}{\alpha_{1i}}\right)\left(e^{-\gamma_i\tau} - 1\right) \\
B_{\phi_{5i}}(\tau) &= -\frac{\alpha_{1i}}{\gamma_i^2}\left(\left(\frac{\alpha_{1i}}{\gamma_i} + \alpha_{0i}\right)\left(e^{-2\gamma_i\tau} - 1\right) + \alpha_{1i}\tau e^{-2\gamma_i\tau}\right) \\
B_{\phi_{6i}}(\tau) &= -\frac{1}{2}\left(\frac{\alpha_{1i}}{\gamma_i}\right)^2\left(e^{-2\gamma_i\tau} - 1\right).
\end{aligned}$$

Consider now a period of length ν 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

$$S(t, T_n) = \frac{1 - P(t, T_n)}{\nu \sum_{j=1}^{n} P\left(t, T_j\right)},$$

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

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

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

$$dS\left(u,T_m,T_n\right)=\sum_{i=1}^{3}\sum_{j=m}^{n}\zeta_j\left(u\right)B_{x_i}\left(T_j-u\right)dW_{iu}^{\mathbb{Q}^{T_m,T_n}},$$

where

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

Notice that, since the $\zeta_j\left(u\right)$ terms are stochastic, the forward swap rates are not normally distributed. We can, however, approximate $\zeta_j\left(u\right)$ 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

$$S\left(T_m,T_n\right)\sim\mathcal{N}\left(S\left(t,T_m,T_n\right),\sigma_N\left(t,T_m,T_n\right)\sqrt{T_m-t}\right),$$

where the volatility σ_N is given by

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

E.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:

$$f_t\left(n,0\right)=\beta_0+\beta_1\exp\left(-n/\tau_1\right)+\beta_2\left(n/\tau_1\right)\exp\left(-n/\tau_1\right)+\beta_3\left(n/\tau_2\right)\exp\left(-n/\tau_2\right). \tag{E.2}$$

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.

E.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:

$$\text{CPR}_\tau = \text{PSA}/100 * \min(.2 * \text{Age}_\tau, 6) \text{ for } t \leq \tau \leq \text{WAM} \tag{E.3}$$

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[1] 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 E.3 to SMMs and compute the implied cash flows as in Section 2. The WAL is then defined as:

$$WAL_t = \frac{\sum_{j=t}^{\text{WAM}} j\, CF_j}{\sum_{j=t}^{\text{WAM}} CF_j}. \tag{E.4}$$

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-

[1]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.

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 for a vintage using nonlinear least squares:

$$\text{SMM}_i^{LR} = b_1 + b_4 \frac{\exp\left(b_2 + b_3 * \text{INC}_i\right)}{1 + \exp\left(b_2 + b_3 \text{INC}_i\right)} \text{ for } i = 1, 2, \ldots, 9 \tag{E.5}$$

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 κ_1 and κ_2 are universal parameters and let χ_t vary across securities and time. To calibrate κ_1 and κ_2, we exploit the fact that as INC $\to \infty$, SMM $\to b_1 + \chi\kappa_1 + (1 - \chi)\,\kappa_2$ (for 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 $\hat{\kappa}_2 = 0.014$. [2]

Given κ_1 and κ_2, there are then four coefficients to be estimated for each security on each date: $\chi_t, b_1, b_2,$ 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.

E.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

[2] We have experimented with alternative calibrations, and obtained qualitatively similar results.

paths of the three factors of the interest rate model using draws of the state variables described in Appendix E.1. We use antithetic variables as a variance reduction technique, giving us 2,000 paths in total.

F 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:

$$P_M = \sum_{k=1}^{T} \frac{X_k(\mathbb{E}r_k)}{\prod_{j=1}^{k} \left(1 + \mathbb{E}r_j + \text{ZVS}\right)}. \tag{F.1}$$

Differently from the OAS, whose formula appears in (2.1), in computing the ZVS both cash 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.[3] Because ZVS abstract from rate uncertainty, following practitioners (for example, Hayre, 2001) we refer to the ZVS-OAS difference as the "option cost":[4]

$$\text{Option cost} \equiv \text{ZVS} - \text{OAS}. \tag{F.2}$$

In Figure A-1, we compare the time series of the ZVS on the TBA index to the OAS.[5] The ZVS (grey 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-1, 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

[3]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.

[4]We note, however, that this is a slight misnomer, as adding volatility to discount rates (in the denominator of equation (F.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.

[5]Both time series and cross-sectional patterns of ZVS and option costs coming from our pricing models have similar properties to those shown here.

of the prepayment option increases in the volatility of the underlying.

Figure A-2 documents the cross-sectional variation in the ZVS and option cost. As shown in 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.[6] 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

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

where σ_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-2: option costs are largest for ATM securities and prices are a concave function of moneyness (and therefore r).

[6]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-1: 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.

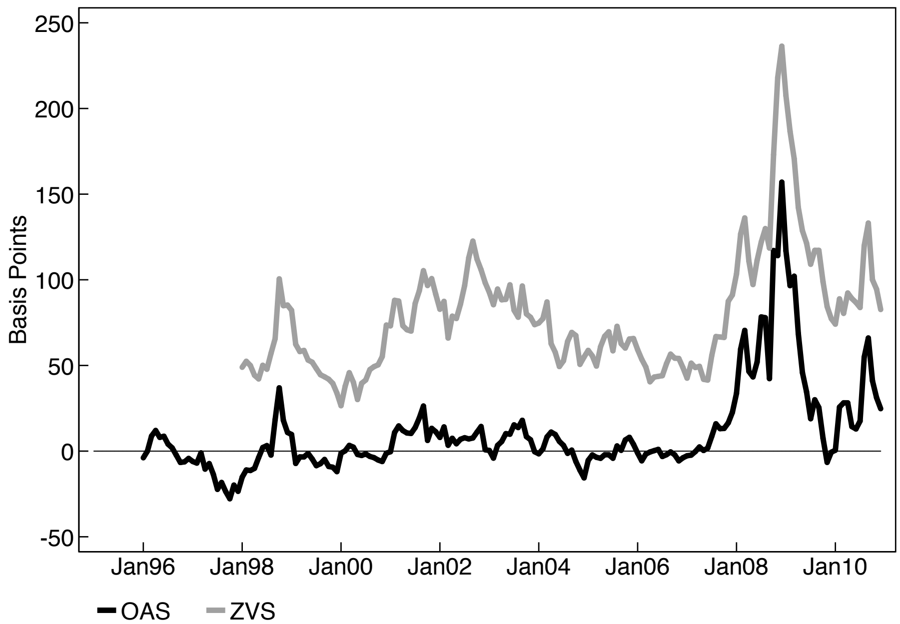
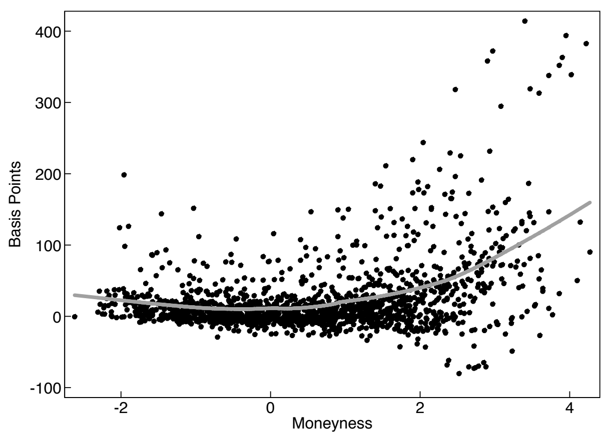


Table A-1: Time-series regressions on TBA index. Coefficient estimates from OLS regression of spreads reported in the top row on the Aaa-Treasury spread, the Baa-Aaa spread, the 10-year to 3-month slope of the Treasury curve and the 2-year into 10-year swaption implied volatility. All regressors are standardized to have zero mean and unit standard deviation. Newey-West standard errors (6 lags) in brackets. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

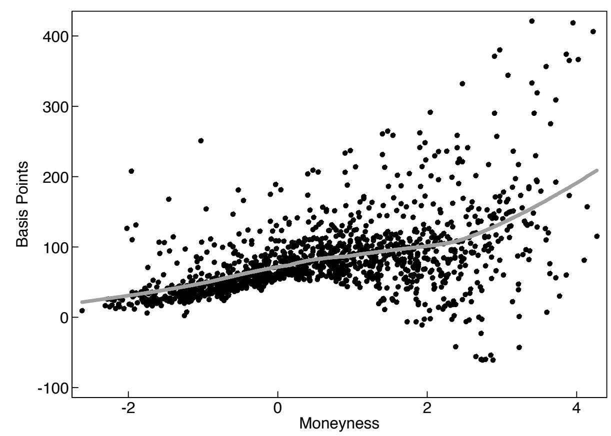
	(1) OAS		(2) ZVS		(3) OAS		(4) ZVS	
Aaa - Treas	3.2**	[1.5]	6.5***	[2.3]	2.3*	[1.2]	6.5***	[2.0]
Baa - Aaa	19.6***	[3.4]	15.5***	[2.2]	9.7**	[3.8]	12.1***	[2.7]
Treas Slope	-0.5	[2.4]	-0.2	[1.3]	2.0	[1.3]	-0.8	[1.5]
Swaption Vol.	2.7	[4.2]	15.5***	[2.5]	-0.8	[2.4]	22.7***	[3.2]
Const	8.8***	[1.8]	70.9***	[1.8]	3.6***	[1.3]	71.8***	[1.5]
Adj. R2	0.76		0.88		0.31		0.77	
Obs.	180		156		139		115	
Dates	199601.201012		199801.201012		199601.200707		199801.200707	

Figure A-2: 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.

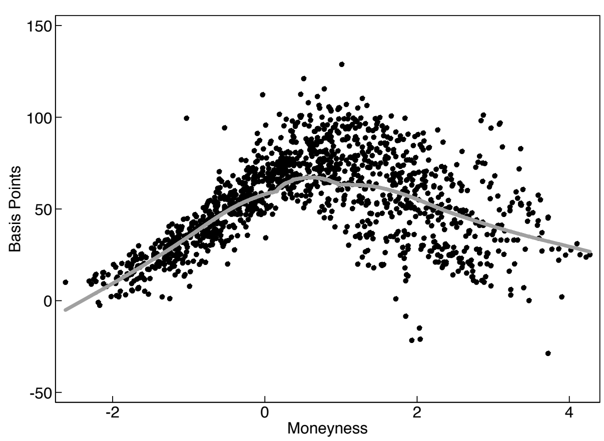
(a) OAS



(b) ZVS



(c) Option Cost



(d) Price