Using Cash Flow Dynamics to Price Thinly Traded Assets

Walter I. Boudry, Cornell University
Crocker H. Liu, Cornell University
Tobias Mühlhofer*, University of Miami
Walter N. Torous, MIT

May 7, 2018

Abstract

This paper investigates the extent to which cash flows as opposed to discount rates are informative in the pricing of thinly traded assets. To do so, we adapt Campbell and Shiller's dynamic Gordon model to the commercial mortgage backed securities (CMBS) market in which, distinct from the stock market, securities are thinly traded but good cash flow data are available. Using our *Self-Propagating Rolling-Window VAR* methodology, we find that cash flows are informative in valuing these thinly traded assets. Our predicted cash flow yields closely resemble ex-post realized transaction yields and even outperform yields based on matrix prices, especially for subordinated tranches. Discount rates, while important, are not as informative as cash flows, except after the financial crisis when the impact of discount rate information increases. We confirm the predominance of cash flow information by directly estimating and not being able to reject the restricted version of our VAR in which CMBS bond cash flow yields are only driven by cash flows. We also show that yields to senior bonds are driven by cash flow growth in subordinate bonds, in line with cash flow information permeating through the CMBS waterfall structure.

Keywords: Thinly traded assets, asset pricing, panel vector autoregression, Commercial Mortgage-Backed Securities.
JEL Codes: G12, R33

*The emails of the authors are: wb242@cornell.edu, chl62@cornell.edu, tmuhlhofer@bus.miami.edu, wtorous@mit.edu respectively. Contact author: Tobias Mühlhofer, School of Business Administration, University of Miami, Department of Finance, 514-S Jenkins Building, Coral Gables, FL 33131. Telephone: +1-305-284-9490.
We thank Caitlin Dillon-Dannhauser, Stefanos Delikouras, Eli Fich, Paul Hanouna, Shawn Howton, Jun Liu, Lorenzo Naranjo, Liang Peng, Rodney Ramcharan, Jay Ritter, Ben Scheick, Paul Tetlock, Rossen Valkanov, Tang Wang, and seminar and conference participants at the Magnolia Finance Conference at Mississippi State University, Shanghai Advanced Institute of Finance, Villanova, the University of Miami, UNC Chapel Hill, and the University of Southern California for helpful comments. We acknowledge the National Association of Insurance Commissioners whose transaction price bond data we use by permission. The NAIC does not endorse any analysis or conclusions based upon the use of its data.

1 Introduction

The relative importance of cash flow versus discount rate information is central to our understanding of the pricing of financial assets. The contributions of these alternative information sources are traditionally evaluated within the discounted cash flow (DCF) pricing framework. Although cash flows generated by an asset should play an important role in it's price formation process, at least in the stock market, variation in the discount factor has been found to play a much more important role.

The stock market is generally characterized by a high level of liquidity. Information about a stock's underlying cash flows, however, is often lacking as dividends are smoothed and are therefore thought to be a poor proxy for the actual cash flows that the firm generates. Interestingly, little is known about the relative importance of cash flows versus discount rates in the informationally opposite situation to common stock in which an asset is thinly traded but reliable cash flow information is readily available. This paper fills this void by investigating what drives security prices in precisely such a setting. In particular, we examine the market for commercial mortgage backed securities (CMBS) which are thinly traded but whose cash flows are transparent and well documented. By doing so, we provide an alternative perspective on the relative importance of cash flow versus discount rate information and so enhance our understanding of the price formation process in financial markets.

Mortgage backed securities are an important asset class and are part of the larger universe of alternative investments created by securitization. These alternative assets tend to possess transparent cash flow information but suffer from thin trading. Aside from the interest the CMBS market has received in the recent literature (see e.g. Titman and Tsyplakov (2010), Childs, Ott and Riddiough (1996), or An, Deng and Gabriel (2011)), in light of the financial crisis, we focus on CMBS to concentrate on the economically more important default decisions of borrowers as opposed to prepayments which, unlike the case of residential mortgage backed securities (RMBS), are effectively prohibited in the CMBS market. Consequently, cash flow volatility in the CMBS market is driven by the default behavior of borrowers. We investigate the CMBS market's drivers of volatility by modeling its price formation process and find that cash flow information does play an important role in this price formation process.

Following the asset pricing literature, our methodology's starting point in modeling the CMBS market's price formation process is the dynamic Gordon model of Campbell and Shiller (1988a), Campbell (1991), and Shiller (1992). This model allows both expected future cash flows as well as expected discount rates to influence an asset's cash flow yield. The challenge in applying the model to thinly traded assets is that cash flow yields are only available when an asset trades. To overcome this obstacle, we develop a rolling window panel VAR methodology, to estimate the cash flow yields of bonds that do not trade as out-of-sample predictions of the VAR. We then use these predicted cash flow yields in the subsequent VAR iteration, joining them with new cash flow data as well as any transaction yields if available, to generate the next set of estimated cash flow yields that we then use in the next VAR iteration. Since each subsequent VAR iteration uses the predictions generated in the previous run, we term this procedure a *self-propagating* rolling window panel VAR. Using this methodology, we estimate close to 75,000 CMBS monthly bond cash flow yields, providing us with a rich panel dataset in which to analyze the determinants of the CMBS bond price formation process.

We first investigate the cash flow yields of a value-weighted portfolio of CMBS bonds and follow Campbell and Shiller (1988a) to decompose the long term forecasts of these cash flow yields into the component related to cash flow growth versus that related to discount rate information. We find that the component of cash flow yields attributable to variation in cash flow growth closely resembles *ex-post* realized yields. This result is consistent with cash flow growth being an important driver of the variation in CMBS bond prices.

We then turn our attention to the panel of individual CMBS bonds and find that the cash flow yields generated by our VAR methodology capture a substantial portion of the variation in actual transaction yields. Using the unrestricted Campbell-Shiller VAR specification, consisting of only cash flow variables, we explain nearly 20% of transaction yields out of sample. In a full specification in which we add discount rate variables, we predict out-of-sample nearly 29% of the variation in transaction yields. Furthermore, our predicted yields outperform yields based on matrix prices in explaining transaction yields. These results are consistent with cash flow information playing an important role in the CMBS market's price formation process. We also find that this conclusion holds both before and after the financial crisis. However, the fraction of variation in transaction yields we predict after the financial crisis when using only cash flows drops to 15%, suggesting a more pronounced role for discount rate information after the financial crisis.

Given the importance of cash flow information in the CMBS market, we then directly test the set of restrictions on the dynamic Gordon model that the corresponding VAR coefficient matrix must satisfy to produce CMBS cash flow yields that are attributable only to cash flow information. Recall that Campbell and Shiller test whether dividend growth drives forecasted dividend yields by estimating the corresponding unrestricted VAR and then testing this restriction on the resultant VAR coefficient matrix using a Wald test. While a Wald test has the advantage that the restricted model does not have to be estimated, we cannot use this technique in our setting because we require that out-of-sample forecasts be forecasted. Instead, distinct from the previous literature, we estimate the restricted VAR directly. We use the restricted coefficient matrix to generate out-of-sample forecasts in our rolling-window setting.[1] We find that the predictive power of the restricted VAR in which CMBS bond cash flow yields are driven only by cash flow information is not statistically significantly different from that generated by the unrestricted VAR. This confirms the predominance of cash flow information in the CMBS bond price formation process.

We also document the informativeness of subordinate CMBS bond cash flows regarding senior CMBS bond yields. Intuitively, senior bond holders observe cash flow shortfalls in the subordinate bonds of their own - or similar deals. This predicts increased risk in the senior bonds themselves and causes senior bond yields to increase reflecting the increased likelihood of a downgrade and increased proximity to default. We find that subordinate bond cash flow growth predicts yields on senior bonds up to a horizon of four months, providing additional evidence of the importance of cash flow information in driving CMBS bond yields. In line with this, we also frequently find that our VAR prices bonds better when estimated jointly over both senior- and subordinate bonds, than for each segment alone.

[1] Estimating the restricted model directly and avoiding a Wald test is also advantageous, in that a Wald test is unstable as it is not invariant to reparameterization. For example, as Campbell and Shiller (1988a) note, the restriction on dividend yields is algebraically equivalent to a restriction on expected returns. However, a Wald test draws different inferences from these two parameterizations.

Campbell and Ammer (1993) were first to investigate the relative importance of cash flow versus discount rate information in fixed income markets. While Campbell and Ammer (1993) study liquid zero-coupon Treasury bonds, by contrast, we concentrate on risky coupon bonds that trade in a thin market, and on decomposing the returns of these bonds. Doing so, we find that cash flows are informative to the price formation process of CMBS bonds. The largest improvement over matrix prices is for non-senior CMBS bonds that have more volatile and therefore more informative cash flows than senior tranches. Consistent with the prior literature on common stock, we also find that discount rates influence CMBS pricing, especially in the post-crisis period, although cash flows exert a relatively stronger impact even then. In contrast to the stock market, most of the variation in CMBS cash yields does not originate from discount rates; yields are primarily driven by cash flow growth in the CMBS market. We further explore whether a distinction in the price formation process exists between the AAA tranches, which are the most liquid bonds, and the rest of the CMBS market and reject this hypothesis.

CMBS differ in their dynamics from corporate debt in significant ways making the CMBS market the preferred setting in which to explore the importance of cash flows in the security price formation process. First, CMBS bonds cannot default. Payments to the bonds are not guaranteed by the sponsor in the same way they are guaranteed by the borrower in a corporate bond setting. This means that CMBS bonds can potentially receive full coupon payments, partial coupon payments, or no coupon payment in a given period. Second, the maturity of CMBS bonds is governed by the flow of principal into the cash flow waterfall structure, rather than by a set maturity date. As such, CMBS bonds may receive cash flows sooner or later than expected. Notice that even though the most senior bonds in the CMBS deal are rated AAA, they face significantly more cash flow volatility than AAA rated corporate bonds. Finally, the senior/subordinate waterfall structure creates a direct link between junior and senior bonds. While the junk bond and investment grade corporate bond universes may be exposed to the same macroeconomic factors, the underlying borrowers differ. In the CMBS market, the underlying mortgages are the same and all that changes between bonds is their exposure to the credit risk of those mortgages. In this way there is a tight information link between the cash flows of junior and senior bonds, which is not a feature of the corporate bond market.

The previous literature has seldom explored the question of the relative importance of cash flow versus discount rate variation in driving asset returns in a panel setting. Vuolteenaho (2002) also relies on a panel setting to investigate the relative importance of cash flow versus discount rate information in driving the returns to individuals assets, in his case, firm-level stock returns. He finds that information about cash flows is the dominant factor driving firm-level stock returns but this highly variable news component is diversified away in aggregate portfolios.[2] This result is consistent with there being more commonality in the discount rate component of stock returns than in the cash flow component. By contrast, we find that cash flow news remains important at the aggregate level in CMBS markets and suggests that CMBS cash flow information is predominately driven by systematic, macroeconomic components. Furthermore, unlike the stock market where a host of economic variables have been identified with apparent in-sample ability to predict returns but fail to deliver consistent out-of-sample forecast gains relative to the historical average (Welch and Goyal (2008)), our rolling VAR procedure more accurately forecasts out-of-sample CMBS yields than the corresponding historical averages by exploiting the systematic nature of CMBS cash flow information.

Although we apply our rolling-VAR methodology to CMBS bonds, it is applicable to any asset that is thinly traded but has consistent cash flow information. Other examples of such asset classes include commercial real estate, natural resource extraction sites (such as mines or oil and gas wells), as well as thinly traded fixed income securities with variable cash flows, such as municipal revenue-based bonds.

The plan of our paper is as follows. Section 2 describes our various data sets in detail, while Section 3 develops our methodology. Section 4 presents our results and Section 5 concludes.

[2]Similarly, Lochstoer and Tetlock (2017) show in a panel setting that cash flow shocks explain more variation in stock market anomaly returns than discount rate shocks.

2 Data

2.1 CMBS Deal Structure

Our sample is made up of bonds from the Commercial Mortgage Backed Securities (CMBS) market. During our sample, the CMBS market was the second largest source of mortgage funding for the commercial real estate market.[3] A typical CMBS deal is based on a senior/subordinate cash flow waterfall. Debt service from the underlying pool of mortgages is passed through to the bonds in sequential order.[4] The principal component of debt service is paid first to the most senior bond class (known as a tranche) until it is fully paid off, then to the next most senior class until it is paid off, and so on. Deals are usually structured in such a way that the senior-most tranches have sufficient credit protection to be rated AAA at issuance, with subordinated tranches having lower ratings. Through time, bond ratings on all tranches can change, according to new information on the likelihood of possible cash flow shortfalls.

There is typically no overcollateralization in the CMBS market, so the face value of the bonds equals the face value of the mortgage pool. The sizing of the senior bond classes is established in such a way as to create bonds that have targeted maturity dates of 3, 5, 7, and 10 years. By volume, the majority of bonds are 10-year bonds because the vast majority of mortgages in the pool typically have a 10-year term.[5]

Interest payments from the mortgage pool are used to pay coupon payments on the bonds. This is once again done in a sequential fashion, with the AAA bonds typically all paid together, and the more junior bonds then paid in sequential order. Most deals have one or more excess interest only classes of bonds (IO bonds). These bonds receive the excess between the weighted average rate paid by the mortgages and the weighted average coupon paid to the bonds.

Losses in a CMBS pool are assigned in reverse order to principal. The most junior bond receives

[3] The mortgages in the CMBS market are senior liens and tend to be made only to stabilized properties. The underlying property types and credit quality is quite diverse and runs the spectrum of the commercial real estate market.

[4] The underlying mortgages in the pool pay debt service on a monthly basis and this is passed through to the bondholders on a monthly basis also. The date of payment is known as the distribution date for the bonds and it is also the date on which remittance reports are released to the market.

[5] This also means that the majority of the bonds look like coupon bonds. Any amortization is paid to the most senior bonds, and all the 10-year bonds are paid back by the balloon payments on the mortgages.

losses first, and once it is wiped out, the next most junior bond starts receiving losses, and so on. This mechanism combined with the sequential ordering of principal and coupon payments is what creates differing levels of credit risk in the bonds. The senior bonds are paid first and receive losses last, while the junior bonds are paid last and receive losses first.

Note that mortgage delinquencies do not necessarily lead to instantaneous cash flow changes to the bonds, even to those bonds most exposed to credit risk. When a mortgage becomes delinquent or is in imminent default, it is transferred to a special servicer. Under the servicing standard, the role of the special servicer is to maximize the net present value (NPV) of the loan. This standard obviously gives the special servicer considerable leeway in dealing with delinquent borrowers. For delinquent loans, the special servicer will advance the debt service the loan should have made, to the extent that they believe it is recoverable. These servicer advances are designed to reduce cash flow volatility in the bonds. Once advances are deemed unrecoverable, the special servicer will no longer advance full principal and interest payments. The process through which this is done is called an ASER (Appraisal Subordinate Entitlement Reduction) and leads to "interest shortfalls" to the most junior bond classes because of the priority ordering of coupon payments.

Unlike in the Residential Mortgage Backed Securities (RMBS) market, prepayment plays less of a role in CMBS cash flow dynamics because of the typical prepayment restrictions placed on the underlying commercial mortgages. Most commercial mortgages will either be locked out from prepaying, require defeasance, or require yield maintenance. Of these mechanisms, only yield maintenance will lead to cash flow volatility in the waterfall. This is not to say there is no prepayment risk in the market, just that it is less significant than in the RMBS market. The risk of receiving principal earlier than anticipated is associated with recoveries received from resolved delinquent loans rather than true prepayments.

2.2 CMBS Data

Our CMBS sample consists of bonds traded by insurance companies during the period 2003 to 2013. Unlike the corporate bond market where transactions are observable through the TRACE system, the CMBS market is a very opaque broker/dealer market. To our knowledge, the only investor group that systematically reports their CMBS trades are insurance companies. These trades are reported in Schedule D Parts 1, 3, and 4 of their regulatory filings. We obtain these filings from the National Association of Insurance Commissioners (NAIC) and SNL Financial. In terms of magnitude, insurance companies are a key purchaser of CMBS, holding approximately 20% to 30% of the market's outstanding bonds at any given time. The institutional landscape is such that insurance companies do not hold the B piece of CMBS transactions, and no secondary market for these bonds exists.[6] Thus, our sample represents the investment grade portion of the CMBS bond market, as well as the portion that is traded (and therefore whose price dynamics it is possible to study).

We calculate the traded price of the bond as the trade size weighted average of the reported dollar prices for that bond on a given day. Because the insurance company filings do not time stamp a trade, it is impossible to identify the last trade of the day. This is the reason why we take the average for that day. The only restriction we place on the price is that the dollar price must be less than 140.[7]

We obtain cash flow information for the bonds from Trepp. Trepp is the leading provider of CMBS market information and the standard source used by market participants to obtain such information. Trepp collects and disseminates information contained in remittance reports. The remittance report contains information on the cash flows flowing into the pool from the mortgages and how this was allocated to each bond class. It also contains information on mortgage delinquencies and also any potential commentary made by the special servicer regarding the delinquent mortgages. Remittance reports are made at the same time that cash flows are distributed to bond holders, so for any given distribution date, the information in the report and the cash flow on the bonds are known to the market at the same time.

[6]The most junior portion of the CMBS deal, which is known as the "B piece" is bought by specialist B piece investors. This usually includes all the bonds up to a BBB- credit rating, and accounts for approximately 5% of the face value of the deal. The B piece is bought through a private transaction at origination and is not usually traded, although the B piece buyer might sell the most senior tranches of the B piece to high yield investors in a one-time transaction. The B piece buyer is given more time and more information (they will have all the information the loan originator has) to do due diligence in making this initial purchase decision.

[7]By convention a dollar price of 100 is par value for the bond. An examination of the data suggests that these are very likely to be data errors and our conversations with CMBS traders also indicate that bonds don't trade at these prices in practice.

From CMAlert's Deal Database we obtain deal and bond level characteristics (bond type, bond credit rating, issuer, subordination level etc.) for each CMBS deal. Using this information we filter our sample to include only fixed rate bonds from fixed rate US CMBS deals.[8] We include only P&I bonds (principal and interest bonds) and exclude all IO bonds.

Our final source of CMBS data is the Thomson Reuters Eikon database. This database provides us with proprietary matrix prices for each bond. Availability of matrix prices is the main limit on our data, as these prices are only available back to 2003. This data, as well as the Trepp cash flow data ends in May 2014. For any tests not requiring transaction prices, our sample ends at this later date.

To create cash flow yields for each bond on a monthly basis, we use the following procedure. For each distribution date, we collect bond cash flow information from Trepp. Because some bonds may be amortizing, we define our cash flow variables as a fraction of the outstanding principal balance. In this sense we don't treat lower coupon payments due to amortization as being informative. For any given month, we include any coupon payment and any prepayment penalties the bond receives as its cash flow for that month. We match the bond cash flow data from Trepp to the most recent matrix price available in the Eikon database, and to the most recent transactions price available from the insurance company filings. For both the matrix price and the transaction price to be valid, we require that it occurred since the last remittance report (i.e. it occurred in the last month.) This ensures that the prices contain information about the latest distribution date.

As is shown in Table 1, in total our sample contains 3981 bonds from 496 CMBS deals and 283,319 bond/month observations. These bonds have an aggregate face value of $652.5 Billion at origination. The earliest deal in the sample is from 1995, while the latest deal is from 2013. Of the 3981 bonds, 2064 have a AAA credit rating at origination and an aggregate face value of $593.2 Billion, 1779 have a non-AAA investment grade rating and an aggregate face value of $56.3 Billion, and 138 have a non-investment grade rating and an aggregate face value of $3 Billion. As discussed earlier, this sample is somewhat biased towards AAA bonds compared to the market as a whole.

The average delinquency rate in the CMBS market has been quite dynamic through time.

[8] WAC bonds are considered fixed rate for our analysis.

Average levels of delinquency were very low in the pre-financial crisis period (less than 2%), and rose quite rapidly during the financial crisis (to over 10%.) Given that the cash flow dynamics in our model are driven largely by delinquencies, this poses the question of whether bonds in the pre-financial crisis period looked just like regular treasuries because of the low average delinquency rates. The answer to this is that while the average delinquency rates were quite low in the pre-financial-crisis period, there was cross-sectional variation in delinquencies. Note that while our estimation starts in 2003, the sample of bonds at this point in time includes seasoned bonds (bonds originated prior to 2003.) The natural ramp up of delinquency in CMBS pools due to seasoning explains why we observe higher pool level delinquencies than average for these seasoned bonds in the pre-financial crisis period.

For our full VAR specification (which includes discount factor variables) we use REIT returns. These are the total returns to the CRSP Ziman REIT Index.[9] For the risk-free rate, we use US-Treasury rates at various maturities. The rate that we term the "Long-Term" interest rate is the interest rate to the 30-year US Treasury Bond where available, and the rate to the 20-year US Treasury Bond where not.

3 Methodology

3.1 Unrestricted VAR Estimation

We focus on the dynamics of the cash flow yields of CMBS bonds to investigate the importance of discount rate versus cash flow information in the CMBS price formation process. A CMBS bond's cash flow yield, analogous to a common stock's dividend yield, expresses the bond's realized cash flow to its current market price.

Cash flow yields are modeled using a dynamic Gordon model (for example, Campbell and Shiller (1988a) or Campbell and Ammer (1993)). We assume that the log cash flow yield, δ, can be expressed as the present value of expected future risk free rates, r, risk premia, π, and cash flow

[9] The returns to the full index and the returns to the Equity-REIT only index have a correlation of .998.

growth rates, Δc:

$$\begin{aligned}\delta_t &= \sum_{j=1}^{\infty} \rho^j E_t[r_{t+j} + \pi_{t+j} - \Delta c_{t+j}] && (1)\\ &= \sum_{j=1}^{\infty} \rho^j E_t[r_{t+j}] + \sum_{j=1}^{\infty} \rho^j E_t[\pi_{t+j}] - \sum_{j=1}^{\infty} \rho^j E_t[\Delta c_{t+j}] && (2)\\ &\equiv \delta_{rt} + \delta_{\pi t} - \delta_{ct} && (3)\end{aligned}$$

where ρ is a log-linearization parameter.[10] Here δ_{rt} is the component of δ_t that forecasts the long-run expectation of the risk free rate while $\delta_{\pi t}$ and δ_{ct} are the components of δ_t that forecast the long-run expectations of the risk premium and the cash flow growth rate, respectively.

Initially, we use an unrestricted VAR model to calculate these multi-period expectations. That is, we forecast the short-run behavior of variables to impute their long-run properties without imposing any restrictions on their behavior. The resultant VAR coefficient matrix provides a quantitative representation of the price formation process and allows us to differentiate the importance of cash flows and discount rate information.

The VAR approach begins by positing a vector of state variables, z_t, to forecast cash flow yields. Our state vector includes the cash flow yield itself, the risk-free rate, and the cash flow growth rate, $z_t = [\delta_t, r_{t-1}, \Delta c_{t-1}]'$. Here δ is the natural log of the ratio of the CMBS bond-level coupon cash flow (interest plus any penalty) divided by the price of the bond (matrix price or transaction price in months in which the bond trades). For r we use the long-term risk-free interest rate, and we use the difference in the log of the coupon cash flow per dollar of outstanding face value for Δc.[11]

Notice that we include cash flow growth rates in the vector z_t but leave out risk premia. In this context we avoid modeling risk premia because CMBS bonds trade in a thin market making it difficult to measure risk premia at any given point in time. On the other hand, we include cash flow growth rates because, unlike common stock, a CMBS bond's cash flows are not managed like

[10]We estimate the value of this parameter as .99 in our dataset.

[11]Defining the cash flow measure in this way prevents unwanted effects inadvertently introduced by the bond's amortization according to a pre-determined schedule, which is also part of the cash flows paid to investors. By contrast, using total cash flows when a bond amortizes would otherwise initially raise the cash flow (as principal is being amortized in addition to interest payments), and then lower it (since the same percentage coupon is paid on less principal outstanding), even though, economically, the bond is still paying its regularly-scheduled coupon service at the pre-determined coupon rate.

stock dividends. In particular, the trustee of a CMBS deal has no discretion in whether to pay a bond or not. Their responsibility is to allocate the proceeds from the underlying mortgages to the bonds according to the rules set out in the cash flow waterfall.

This implies that we obtain components δ_{rt} and δ_{ct} by forecasting risk free rates and cash flow growth rates, respectively, directly from the VAR and attributing any remaining component of δ to risk premia. In particular, we assume that z_t can be written as

$$z_t = Az_{t-1} + u_t \tag{4}$$

or

$$\begin{bmatrix} \delta_t \\ r_{t-1} \\ \Delta c_{t-1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} \delta_{t-1} \\ r_{t-2} \\ \Delta c_{t-2} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix} \tag{5}$$

and

$$E_t[z_{t+k}] = A^k z_t. \tag{6}$$

Equation (6) indicates that, once estimated, the VAR coefficient matrix, A, can be used to make predictions of the state vector z, by multiplying the vector of time-t realizations of the state variables by the coefficient matrix once for each forward period of predictions one wishes to make. Campbell and Shiller provide a closed-form solution for the infinite-horizon predictions, and show how these can be decomposed into components of predicted yield, which are driven by each individual state variable. It follows that if $\iota 2 = [0\ 1\ 0]'$ and $\iota 3 = [0\ 0\ 1]'$, we have that

$$\delta_{rt} = \sum_{j=1}^{\infty} \rho^j \iota 2' A^j z_t \tag{7}$$

$$= \iota 2' A(I - \rho A)^{-1} z_t \tag{8}$$

and

$$\delta_{ct} = \sum_{j=1}^{\infty} \rho^j \iota 3' A^j z_t \tag{9}$$

$$= \iota 3' A (I - \rho A)^{-1} z_t. \tag{10}$$

Under an assumption of constant expected returns, CMBS cash flow yield can be modeled as:

$$\delta_t^* = \iota 2' A (I - \rho A)^{-1} z_t - \iota 3' A (I - \rho A)^{-1} z_t \tag{11}$$

The long-run expectation of the CMBS bond's risk premium is then captured by the difference between the observed, δ, and modeled, δ^*, cash flow yields.

We can then compare realized CMBS cash flow yields, δ_t, to δ_{ct} and δ_{rt}. If, for example, realized yields resemble δ_{ct} this then would indicate that cash flow growth information is an important driver of CMBS bond prices.

Using this framework, we can also forecast, say, one-period ahead CMBS cash flow yields, δ_{t+1}. Given observable underlying cash flows, c_{t+1}, one can then predict the CMBS bond price, P^*_{t+1} that can be used to value the bond even in the absence of transactions in the bond.

3.2 Examination in a Panel Setting

Since we are interested in understanding the yield dynamics of individual bonds, we also focus on a bond-by-bond level empirical analysis. This implies we are dealing with panel data consisting of a cross-section of individual bonds over time and so we adapt the VAR approach to this setting. In a panel setting, we stack time series observations for individual bonds. We then modify Equation (5) as follows:

$$\begin{bmatrix} \delta_{i,t} \\ r_t \\ \Delta c_{i,t} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \delta_{i,t-1} \\ r_{t-1} \\ \Delta c_{i,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,i,t} \\ u_{2,i,t} \\ u_{3,i,t} \end{bmatrix}. \tag{12}$$

The above representation shows that we model the yield for bond i at time t, through lagged values of the yield for the same bond, the common long-term interest rate (which is identical for all bonds at a particular point in time), and the lagged cash flow growth for that bond. The coefficient matrix estimated in this way characterizes the common price formation process applied across all bonds in the sample used, with the cross-section giving additional statistical power in the estimation of the coefficients. [12]

However, we face an additional challenge in using this empirical framework in a thinly traded market. In particular, in order to estimate a VAR coefficient matrix to model a bond's cash flow yield requires a full set of time series observations on the yield itself. But since a yield requires a price, and a price requires a transaction, a consequence of working in a thinly traded market is that we cannot construct a full time series of the cash flow yield itself. We overcome this problem by taking advantage of our empirical framework in the following way: once a VAR coefficient matrix is estimated, we apply this coefficient matrix to existing information to infer the cash flow yield of a bond in between trades as an out-of-sample VAR prediction. This becomes our *self-propagating* rolling VAR, whose mechanics are explained next.

3.2.1 Self-Propagating Rolling-Window VAR Procedure

The CMBS bond market is too thinly traded to estimate a VAR such as Equation 12 solely using yields based on transaction prices. Instead, we use a rolling-window VAR to populate the data

[12]It is well known that in panel data studies, unobserved systematic cross-sectional heterogeneity can lead to biased OLS standard errors and thus to incorrect statistical inferences. For this reason, such studies often use fixed effects and clustering of standard errors at various levels of panel observation (see e.g. Love and Zicchino (2006), who also do this in their panel VAR study, although they have a different motivation for doing so, as they have economic interest in these coefficients). We elect not to incorporate such techniques in the procedures we use. This is because, while we do present VAR coefficients and hypothesis tests associated with them in our study, these purely serve to convey to the reader the basic interrelationships between the state variables involved in our study, and are not important for our primary result. Our primary interest lies in the out-of-sample predictions from the panel VARs we estimate, which should not be biased by this omission. On the contrary, it has been shown (for example, Wooldridge (2010)), that in the presence of unobserved systematic cross-sectional heterogeneity in panel data, OLS coefficients may become inefficient. If this is the case, this should negatively impact – rather than overstate – the predictive power of our panel VARs. While Petersen (2009) argues in this context that the use of random-effects estimation would help this situation, we elect to omit this for two reasons. First, as far as we know, the statistical properties of a GLS-based random-effects estimator in the context of a panel VAR are not sufficiently well understood. Second, a meaningful run of such a procedure would require a large amount of data, both in the cross-section and the time series. Therefore, this would preclude us from being as parsimonious in our approach as we currently are, which would force us to exclude a large part of the early portion of the data sample.

with VAR-predicted yields where a transaction yield does not exist. The primary difficulty in this context becomes estimating the VAR system which is then used to predict yields, vis-a-vis this lack of data. Specifically, we proceed as follows.

Consider an unbalanced panel dataset for a set of CMBS bonds, $i \in 1, 2, 3, \ldots, I$, over time, $t \in 1, 2, 3, \ldots, T$. We are tasked with populating cash flow yield values in this panel, which are missing due to the absence of trades in a bond. We generate a *mixed yield* for each bond, i, at time t, which we initially populate (where available) with a matched transaction yield. That means, for each bond, i, in whose market segment, m, one or more transactions occur at time t, we set $mixed.yield_{i,t}$ equal to the average of transaction yields in market segment m at that time. We define a bond's matched market segment m, by the interaction of its vintage (i.e. the year in which the deal was issued), and its initial credit rating (which proxies for the bond's level of subordination within its deal). A matched market yield defined in this way provides a reasonable proxy for a bond's own yield in the absence of trades. Applying this procedure fills in yields for some panel observations but still leaves the vast majority of yield observations blank. In practice, out of 17,290 month-vintage-rating combinations, only 4,077, or 23.6% of matched segments have any transactions and are therefore populated this way. It will be the task of our Self-Propagating Rolling VAR to fill in the rest.

We specify a time window, w, over which to run each iteration of the rolling-window VAR. For the initial VAR run, which uses observations for all bonds present in the dataset from $t = 1$ to $t = w$, we have no choice but to fill in the missing yields using matrix prices, since we require continuous data to estimate the VAR system. After estimating the coefficient matrix for the initial VAR, however, we then fill in all unpopulated yields (i.e. yields for CMBS bonds for which no matched yields were available) for time $t = w + 1$ as predictions from the panel VAR just estimated. The *mixed yield* therefore becomes a mix of matched yields (where available) and predicted yields where not.

We then move the estimation window up by one period and re-estimate the VAR from $t = 2$ to $t = w + 1$. It should be noted that the previously unpopulated yield observations at $t = w + 1$ are now populated with predictions from the previous VAR, which are then matched up with new data for the other state variables (cash flows and interest rates), all of which are used in the estimation. The one-period forward predictions from this newly estimated VAR are then used to populate the previously unpopulated yield observations for $t = w + 2$. These are then combined with new data for the other state variables in that period to be used for the VAR estimated from $t = 3$ to $t = w+2$, to generate predicted yields for $t = w + 3$, and so forth. Since the VAR, in part, generates its own data, we term it *self-propagating.*

It should be apparent that due to the structure of the panel VAR, we generate a specific new yield for each individual bond simultaneously, rather than an overall market yield. To see how this works, consider a panel of four bonds ($i = 1, \ldots, 4$), for which we have estimated the coefficient matrix, as illustrated in Equation 12. The bond-specific predictions of the state variables are then calculated by multiplying the time-t matrix of bond-by-bond values of state variable realizations with the transpose of the coefficient matrix[13], as follows:

$$\begin{bmatrix} \delta_{1,t} & r_t & \Delta c_{1,t} \\ \delta_{2,t} & r_t & \Delta c_{2,t} \\ \delta_{3,t} & r_t & \Delta c_{3,t} \\ \delta_{4,t} & r_t & \Delta c_{4,t} \end{bmatrix} \begin{bmatrix} a_{1,\delta} & a_{1,r} & a_{1,\Delta c} \\ a_{2,\delta} & a_{2,r} & a_{2,\Delta c} \\ a_{3,\delta} & a_{3,r} & a_{3,\Delta c} \end{bmatrix} = \begin{bmatrix} \delta_{1,t+1} & r_{t+1} + \epsilon_1 & \Delta c_{1,t+1} \\ \delta_{2,t+1} & r_{t+1} + \epsilon_2 & \Delta c_{2,t+1} \\ \delta_{3,t+1} & r_{t+1} + \epsilon_3 & \Delta c_{3,t+1} \\ \delta_{4,t+1} & r_{t+1} + \epsilon_4 & \Delta c_{4,t+1} \end{bmatrix} \tag{13}$$

Of primary interest in this case are the predictions for δ for each bond. As described above, we then take these yields and fill them in for bonds at time $t + 1$ where this field is unpopulated (i.e. no matched yield exists). Before doing this, of course, they must be re-meaned and exponentiated, as the VAR is run with demeaned logs of variable realizations.

There may be a concern that using previous predictions in a new estimation run may cause the new predictions (which would cumulate errors) to become extremely noisy, or otherwise to slowly decay to zero, like an impulse-response function. Our results show, however, that this is not the case. There are two reasons for this. First, in the panel, some new yield data are used, since matched transactions and some trades occur, which supplements the information in the predictions.

[13] In this coefficient matrix, the second subscript refers to the equation from which the coefficient came, so, for example $a_{2,\delta}$ refers to the second coefficient (i.e. the one on interest rate) from the δ equation.

Second, all other state variables contain new information which is processed in the new estimation run.

We choose to include bonds that, for a given number of lags L included in the VAR, have at least $2L + 1$ time-series observations. This restriction helps ensure that enough of a time series exists to reliably estimate autoregressive coefficients, even within bonds and not just across them. Further, for bonds that first appear in the dataset at a time beyond the initial window from $t = 1$ to $t = w$ (where we automatically filled unpopulated *mixed.yield* observations with matrix-price based yields), we populate any missing *mixed.yield* observations in the bond's first $2L+1$ months of existence with matrix price based yields. Otherwise, if we did not do this, we could only use bonds for our estimation which in their first $2L + 1$ months had matched transactions in their market. This would be a non-randomly drawn sample and would therefore introduce selection bias. Overall, the worry that we fill in much of our data this way, leaving no room for our self-propagating VAR, would be unfounded: in practice, our self-propagating VAR is left with the task of estimating 74,708 yields. To be parsimonious, and to use as few matrix prices as possible, we use a short window length of $w = 36$ months. With much shorter window sizes, estimates become exceedingly noisy. Further, it may be economically justifiable to estimate the VAR over a sample that covers a substantial part of a market cycle. It should be noted here that our rolling window procedure allows the VAR coefficients to vary over time. Again this should be economically justified, as investors may make different use of the underlying data under different economic scenaria.

Besides estimation over the entire dataset, we also conduct this procedure on two subsets of the CMBS market separately: bonds rated AAA at issuance and bonds that are not. This allows us to distinguish between bonds of senior tranches and those of subordinate tranches. Economically, it is conceivable that the price formation process (and therefore the VAR coefficients) for the different market segments may differ.

3.2.2 Measuring Predictive Performance

We construct a variety of statistics to examine the predictive properties of our self-propagating VAR. We first use the ratio of the standard deviation of the yields we predict over the standard deviation of a realized (or matrix price based) yield series. However, since we predict out-of-sample, if we were to produce VARs that were extremely noisy but whose estimated yields in no way resembled ex-post realized yields, this measure would still be high (and could even be greater than one). Clearly this would not be indicative of a high-quality yield forecast. Therefore, we augment this measure by also reporting the correlation coefficient between the predicted and the realized (or matrix price based) yields.

We begin by comparing our predicted bond yields to yields implied by matrix prices. This allows us to gauge whether our VAR produces results that are statistically reasonable. While matrix prices themselves are not always indicative of transaction prices, this comparison is still useful because it allows us to assess our results over the entire sample of yields produced. In contrast, the comparisons that follow use only yields produced around bond trades.

Next, we assess our self-propagating VAR's ability to predict eventual transaction yields (when the bond is sold). Since these are out-of-sample predictions of a value that is realized ex-post, we also introduce the additional measure of the out-of-sample R-squared (see Welch and Goyal (2008)) defined as

$$R^2_{OOS} = 1 - \frac{\sum_{t=w}^{T}(\delta'_{t+1} - \delta_{t+1})^2}{\sum_{t=w}^{T}(\overline{\delta_t} - \delta_{t+1})^2} \qquad (14)$$

In this expression δ_{t+1} is the ex-post realized transaction yield at time $t+1$, δ'_{t+1} is the predicted yield and $\overline{\delta_t}$ is the historical average yield over the rolling window ending at time t. This measure compares the sum-of-squared prediction errors to the prediction error obtained by using the historical mean as the best predicted yield. The more of an improvement the prediction offers over the historical mean, the closer to one this statistic becomes. If the prediction does no better than the historical mean, the statistic is zero; if it does worse, the statistic is negative.

Next, we compare our *mixed.yield* with transaction-based yields. Since mixed yields are augmented by transactions that occur in other market segments at a particular time, we compare, for every trade at time t, the prior month's mixed yield (i.e. time $t-1$), to ensure that we are examining a prediction, and not a value that potentially contains the information of the subject trade itself. Here, too, we report the ratio of standard deviations, the correlation coefficient, and the out-of-sample R-squared.

Lastly, we benchmark our predicted yields against matrix price based yields. To do this, we report the same statistics as for $mixed.yield_{t-1}$, for the matrix price at time $t-1$. Once again, we use the previous month's price since, if we were to report the matrix price at time t, this could potentially contain information from the subject trade, and therefore would not be a prediction.

3.3 Restricted VAR Estimation

To better understand the relative roles of cash flow versus discount rate information in the context of the dynamic Gordon model, Campbell and Shiller (1988a), Campbell and Shiller (1988b) and Shiller (1992) specify a set of restrictions that their corresponding VAR coefficient matrix must satisfy to produce yields that are attributable only to cash flow growth information.

These restrictions give rise to the infinite-horizon forecasts, δ_{ct} and δ_{rt}, which we construct in our aggregate-level analysis. In our rolling VAR setting, we produce one-period forward forecast yields and so the problem takes on a different form. Specifically, over a one-period forecast horizon, the VAR coefficient matrix must fulfill the following set of restrictions:

$$\iota 1'(I - \rho A) - (\iota 3' - \iota 2')A = 0. \tag{15}$$

A VAR coefficient matrix that conforms to this set of restrictions will produce one-period forward forecast yields that are driven only by cash flow growth information.

Campbell and Shiller define these restrictions for a single-lag VAR. By contrast, in our case we use a four-lag VAR. The restrictions for a VAR that contains more than one lag take on the same functional form, as long as the coefficient matrix A is in *companion form* and the selection vectors $\iota_1, \iota_2, \iota_3$ are suitably adjusted to skip trivial (i.e. deterministic) coefficients.

Specifically, our four-lag VAR in companion form becomes:[14]

[14]While in the rest of our discussion we use r as the second state variable and Δc as the third to correspond to the common $r - g$ form of the Gordon Growth Model, Campbell and Shiller, in their formulation of the restrictions put Δc second and r third. The order, of course, matters only for consistency with the selector variables ι_n, and so we

$$
\begin{bmatrix} \delta_t \\ \delta_{t-1} \\ \delta_{t-2} \\ \delta_{t-3} \\ \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ r_t \\ r_{t-1} \\ r_{t-2} \\ r_{t-3} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} & a_{1,6} & a_{1,7} & a_{1,8} & a_{1,9} & a_{1,10} & a_{1,11} & a_{1,12} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} & a_{2,6} & a_{2,7} & a_{2,8} & a_{2,9} & a_{2,10} & a_{2,11} & a_{2,12} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} & a_{3,6} & a_{3,7} & a_{3,8} & a_{3,9} & a_{3,10} & a_{3,11} & a_{3,12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \delta_{t-1} \\ \delta_{t-2} \\ \delta_{t-3} \\ \delta_{t-4} \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ \Delta c_{t-4} \\ r_{t-1} \\ r_{t-2} \\ r_{t-3} \\ r_{t-4} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ 0 \\ 0 \\ 0 \\ u_{2,t} \\ 0 \\ 0 \\ 0 \\ u_{3,t} \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (16)
$$

The selection vectors become $\iota 1 = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]'$, $\iota 2 = [0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0]'$, and $\iota 3 = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0]'$. With the coefficient matrix and the selection vectors defined this way, the reduced-form expression of the lefthand side of equation (15) implies the following 12 cross-equation restrictions:

adopt their order in this section in order to not have to re-specify the ι_n variables.

$$
\begin{bmatrix}
\rho a_{1,1} + a_{2,1} - a_{3,1} + 1 \\
\rho a_{1,2} + a_{2,2} - a_{3,2} \\
\rho a_{1,3} + a_{2,3} - a_{3,3} \\
\rho a_{1,4} + a_{2,4} - a_{3,4} \\
\rho a_{1,5} + a_{2,5} - a_{3,5} \\
\rho a_{1,6} + a_{2,6} - a_{3,6} \\
\rho a_{1,7} + a_{2,7} - a_{3,7} \\
\rho a_{1,8} + a_{2,8} - a_{3,8} \\
\rho a_{1,9} + a_{2,9} - a_{3,9} \\
\rho a_{1,10} + a_{2,10} - a_{3,10} \\
\rho a_{1,11} + a_{2,11} - a_{3,11} \\
\rho a_{1,12} + a_{2,12} - a_{3,12}
\end{bmatrix} = \underline{0}, \tag{17}
$$

where $\underline{0}$ is a zero vector of length 12.

Campbell and Shiller test whether cash flow growth drives one period forecasted dividend yields by estimating their unrestricted VAR and then testing the restrictions on the resultant VAR coefficient matrix using a Wald test. The literature has tended to follow Campbell and Shiller and rely on the Wald test,[15] which has the great advantage that the restricted model does not have to be estimated. In our case, however, this approach is not feasible since we generate out-of-sample forecasts. If we generated these from an unrestricted VAR, this would not isolate cash flow information as a driver of these yields. If we wish to test the importance of cash flow information in this setting, we have no choice but to estimate the restricted VAR and use the restricted coefficient matrix to generate our predicted yields. Having done this, we can then compare the prediction error from the unrestricted rolling VAR to that from the restricted rolling VAR. If the two sets of prediction errors are statistically indistinguishable, this means that the restrictions did not significantly *handicap* our VAR, indicating that cash flow growth information constitutes

[15] For an exception, see Mühlhofer and Ukhov (2011) who also apply this restricted VAR estimation procedure.

the primary driver for yields.

Estimating the restricted VAR has an additional advantage. It is generally known that the Wald test is unstable as it is not invariant to re-parameterizations of a set of restrictions. For example, as noted by Campbell and Shiller, the restriction on the dividend yield is algebraically equivalent to a restriction on expected returns and yet a Wald test yields different inferences from the two parameterizations. Testing the restrictions using actual estimation of the restricted model helps overcome this difficulty.

Estimation of restricted VARs is not common in the literature, and so we explain our procedure here. While it is technically possible to construct the likelihood function for the restricted VAR and to maximize this likelihood explicitly using a numeric optimization engine, doing so is computationally intensive (especially with the size of dataset that we have in this study). Instead, to estimate the restricted VAR, we use a technique that can be generally applied to simultaneously estimated systems of equations, of which the restricted VAR is a special case. A baseline result exists which shows that for a system of nested models equation-by-equation OLS is the most efficient estimation method. However, a VAR, when restricted, is no longer a system of nested models and so simultaneous estimation is necessary.[16]

In particular, assume that a set of restrictions on a system of equations is specified as follows:

$$R\beta^R = q \tag{18}$$

where β^R is the vector of restricted coefficients from all equations, and R and q, respectively, are a matrix and a vector which specify the left-hand and right-hand sides of a set of linear restrictions. The estimator for the restricted coefficients becomes the solution to the Lagrangian that minimizes squared residuals subject to the restrictions, or:

[16] This estimation technique is described in textbooks like Greene (2003), or more concisely in Henningsen and Hamann (2007), whose notation we borrow for this exposition, and whose R library we use to conduct the estimation.

$$\begin{bmatrix} \widehat{\beta}^R \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X^T\widehat{\Omega}^{-1}X & R^T \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} X^T\widehat{\Omega}^{-1}y \\ q \end{bmatrix} \tag{19}$$

where λ is a vector of Lagrange multipliers of the restrictions, and $\widehat{\Omega}$ is the matrix that characterizes the covariance structure of the disturbance terms, customary in GLS or feasible-GLS estimators. The first term on the right-hand side in the equation above also constitutes an estimator for the covariance of the estimated coefficients.[17]

For our restricted-VAR estimation, we apply the technique above for each iteration of the *self-propagating* rolling VAR, to produce a set of yields calculated from a restricted coefficient matrix (and therefore driven by cash flow information only). As before, we then assess the performance of the model, by comparing predicted yields to ex-post realized yields when a CMBS bond trades. Finally, in order to compare the performance of the restricted VAR procedure to that of the unrestricted VAR procedure, we conduct a Diebold-Mariano test with Harvey, Leybourne, Newbold modification (see Diebold and Mariano (1995), and Harvey, Leybourne and Newbold (1997)) on the prediction errors from the two models. If the prediction errors from the restricted VAR are not statistically distinguishable from those from the unrestricted VAR, this means that the VAR naturally follows the restrictions and that cash flow growth drives yields.

While the direct estimation of the restricted Campbell-Shiller VAR is strictly necessary in our setting, by illustrating a computationally feasible way to do this, we also open avenues for further research in the extensive literature that relies on the dynamic Gordon model.

4 Results

4.1 Value-Weighted CMBS portfolios

To begin with, Table 2 gives the results of decomposing the cash flow yields of a value-weighted portfolio of our sampled CMBS bonds. We also present these results separately for the CMBS

[17] Mühlhofer and Ukhov (2011) compare results obtained this way to results obtained by numerically maximizing the likelihood for the restricted VAR and find the two methods to yield the same estimates.

bonds rated AAA at origination versus the CMBS bonds not AAA rated at origination.

The correlation between observed yields, δ_t, and the cash flow component, δ_{ct}, is 0.820 over the entire sample period from January 2003 to May 2014. The correlation is higher for non-AAA rated bonds (0.8999) than AAA rated bonds (0.6481). All of these correlations are statistically significantly different from zero. In other words, cash flow expectations play an important role in determining CMBS yields. This is especially true for lower rated bonds whose position towards the bottom of a securitization's waterfall makes their pricing particularly sensitive to the cash flow prospects of the underlying properties.

The informativeness of cash flow expectations increases after the financial crisis, January 2009 to May 2014, than before the financial crisis, January 2003 to December 2007. For example, if we rely on the VAR estimated over the entire sample period but correlate separately yields and cash flow expectations before and after the financial crisis, the correlation is .9256 after the financial crisis but only 0.5476 before the financial crisis. Again, cash flow information is more important for non-AAA rated bonds than AAA rated bonds. The same conclusions hold if, alternatively, we separately estimate VARs before the financial crisis and after the financial crisis. Now the correlation between CMBS yields and cash flow expectations is 0.9926 after the financial crisis as compared to only 0.3965 before the financial crisis. Intuitively, cash flow information is more important for CMBS bonds after the financial crisis because the significant decline in the values of commercial properties during the financial crisis left CMBS bond values particularly sensitive to the cash flows and values of the underlying properties.

While cash flow expectations are important in determining CMBS yields, especially for non-AAA rated bonds after the financial crisis, interest rate expectations do not play as important a role. In unreported results, we find that the correlation between observed yields and the interest rate component, δ_{rt}, is only 0.0501 over the entire sample period and is not statistically significant. This correlation for non-AAA rated bonds is 0.0776. Furthermore, the lack of informativeness in interest rate expectations is also evidenced by the fact that the correlation between AAA yields and δ_{rt} is -0.4026 which while statistically significant is the opposite sign of what we would expect in the dynamic Gordon model.

In summary, the results of Table 2 confirm the importance of cash flow information in the pricing of CMBS bonds in the aggregate. In the next section, we test this question for individual bonds. In order to do so, we exploit this informativeness and investigate how to use cash flow information to forecast CMBS yields.[18]

4.2 Individual CMBS bonds: Unrestricted VAR

To establish the feasibility of a panel VAR in this setting, and to investigate its behavior in a more controlled environment, we begin by running a single panel VAR over our full sample. For this estimation only, we fill yields for untraded bonds from matrix prices. Later, in the rolling setting, these will then be filled by our *self-propagating VAR*.

We begin by examining the coefficient matrix from our full sample panel VAR shown in Table 3. The state variables used are *yield*, the log of the cash flow yield, $lt.rate$, the log of the long-term risk-free interest rate, and $cf.growth$, the change in monthly log cash flow. The table shows the coefficients for the VAR using one lag. Although we run these VARs for up to four lags, we omit tabulating the coefficients for our higher-lag VARs to save space. The VARs are estimated through equation-by-equation OLS.

Notice in the first line of the coefficient matrix that a very large portion of the variation in the yield (88 percent) is accounted for by the combination of a one period lag in the yield, long term interest rate, and cash flow growth.[19] The t-statistics are large in this case, in part due to our large number of observations, so we do not place great emphasis on these hypothesis tests. We highlight, rather, the signs of the coefficients. These are positive for the lagged yield (i.e. the yield is highly persistent), positive for the risk-free rate, and negative for cash flow growth. Since the interest rate and cash flow growth represent the denominator $(r - g)$ in the Gordon growth model, this suggests that our VAR does a good job in simulating this framework.[20]

[18] For robustness, we have estimated these results with longer lag structures in the VAR, and different time horizons for the risk-free rate. The results are qualitatively maintained and are available upon request.

[19] In untabulated results, we find that a two-lag VAR model does not offer much additional information over a one lag VAR model in terms of each equation's R^2. This seems to suggest that the market is fairly quick at integrating this information into prices. One caveat here, of course, is that the majority of yields comes from matrix prices, and so this statement may apply more to the algorithm used to construct these, rather than to the market's actual price formation process.

[20] Recall that both r and g are linear since we are using logs.

Table 4 presents statistics comparing the predicted yields from our single panel VAR, with realized yields where available, and otherwise matrix price based yields. Specifically, the table presents, for each VAR specification, the ratio of the standard deviations between the predicted and the realized or matrix price based yields. All predicted yields are fitted values from a set of full-sample panel VARs, with one, two, three, and four lags. The first line shows yields from a VAR system that contains only the Campbell-Shiller variables (yield, the risk-free rate, and cash flow growth). The second line shows yields from a system that, in addition, contains two additional discount rate variables (squared cash flow growth and REIT returns), which we further explore later in a rolling framework, as an alternative specification.

It is apparent from Table 4 that, in sample, our VAR systems capture a substantial amount of the variation of CMBS yields. In fact, all ratios of standard deviations are at or above .94. The first line of the table indicates that cash flow information by itself already captures a large portion of the variation of CMBS yields. Comparing these values with those in the second line leads to the conclusion that additional discount rate information to the VAR gives very little improvement in the amount of variation captured. The same can be said about increasing the number of lags in the system: little incremental variation of yields is captured by the VAR in this way. It is important to view this result with caution, as in this setup the majority of the yields used consists of matrix price based yields.

We now turn to the results illustrating the performance of the *Self-Propagating* Rolling-Window Panel VAR estimation of yields for CMBS bonds. Table 5 presents these results, for the unrestricted cashflow-only specification. Panel A shows the results for the full time series. This means that, given that our data starts in 2003 and accounting for the lags needed in the VAR as well as the length of our estimation window, the first prediction is for the beginning of 2006. The first section of the panel shows how our model performs for all bonds, while the subsequent sections provide results for estimations conducted over only the set of bonds rated AAA at origination (i.e. senior securities), and only the set of bonds rated non-AAA (i.e. subordinated securities). We conduct separate estimation runs for each of the three segments, in order to allow the VAR coefficient matrices to vary among market segments.

The first column ($Pred_t$ vs Mat_t) compares our predicted yields to yields based on contemporaneous matrix prices. We report the ratio of the standard deviation of our predicted series over that of the matrix-price based series, to show what fraction of the variation we capture. Since we are making out-of-sample predictions, in order to distinguish between a set of noisy predictions and one that actually captures the variation in matrix prices, we also report the correlation between the two series and a t-test, testing the null hypothesis of zero correlation. For all bonds, we find that we capture .3406 of the variation, with a positive significant correlation coefficient of .1788. It should be remembered that matrix prices are not transaction prices. Therefore, matching these as closely as possible is not our primary goal. Rather, we rely on these estimates to better understand the properties of our estimation procedure. In particular, the fear that re-using previously generated predictions in subsequent iterations of the VAR would lead to extreme noise in our predictions (or perhaps a gradual dying off in their variance) is seen to be unfounded. Performing this comparison with matrix prices for this purpose has the added advantage that we have observations for nearly the entire panel, while transactions tend to be scarce.

The second column ($Pred_t$ vs $Trans_t$) compares our predicted yields to actual realized transaction yields. The ratio of standard deviations is .9112 and the correlation coefficient is positive and significant at .5634. The Out-of-Sample R-squared (R^2_{OOS}) is .1964 and indicates that our VAR estimation procedure based on a dynamic Gordon growth model captures a substantial amount of the variation of actual CMBS bond transaction yields.[21] Recall that for much of our VAR estimation here we do not have actual yields but rather use predicted yields from our VAR instead in our *self-propagating* procedure. This suggests that in a thinly traded market in which good cash flow information is available, the dynamic Gordon growth model is an effective way to characterize the market's price formation process.

The third and fourth columns ($Mixed_{t-1}$ vs $Trans_t$ and Mat_{t-1} vs $Trans_t$), respectively, compare the $mixed.yield$ the period before a transaction, as well as the matrix price based yield the period before a transaction, to the actual transaction yield. Viewed jointly, these compare the quality of our VAR predictions with matrix prices. The third column of the table shows a

[21]By comparison, R^2_{OOS} below 10% are very common in the stock market predictability literature.

correlation coefficient that is similar to that in the second column (.5548 versus .5634), with a similar ratio of standard deviations (.9599 versus .9112), and a lower R^2_{OOS} (.1964 versus .1515). This indicates that in this context our estimates are overall noisier than the ones reported in column 2, which makes sense as the estimate comes from one month before the trade. However, we still capture a high fraction of yield variation, especially when compared to matrix prices. For these (in column 4), there is a positive significant correlation of .246, with a ratio of standard deviations of 1.89 and an R^2_{OOS} of -2.58. This indicates that these matrix prices are extremely noisy, while our VAR produces much more stable estimates.[22]

The second section of Panel A shows results for senior bonds only (i.e. bonds that were rated AAA at origination). The first column indicates that here, too, we produce statistically reasonable estimates. The second column shows that we predict a slightly higher fraction of the variation of transaction yields with the R^2_{OOS}, as the most telling statistic, at .2254 instead of .1964. This is likely due to the fact that information in the senior bond market is more plentiful, and that overall these bonds exhibit lower price volatility and are therefore easier to price.

The third column for senior bonds also reveals that, according to the R^2_{OOS}, our mixed yields the period before a trade predict transaction yields better in this subsample than they do in the entire sample. Not only this, but mixed yields the period before a trade are actually a slightly better predictor of transaction yields than pure predictions during the period of the trade (.2254 versus .2322). This is likely due to the matched yields that we use to form mixed yields. In the senior bond segment, transactions are plentiful and so are matched yields. Further, there should be less heterogeneity among bonds in this universe than in the entire sample, and so the matched yields that are substituted are probably a closer proxy for the yield of the bond in question. The fourth column, once again, shows that in this subset, matrix prices, though slightly better than before, are still noisy, with a ratio of standard deviations of 1.23 and an R^2_{OOS} of $-.55$. Here, too, our VAR predictions seem to be more stable.

The final section of Panel A shows the results for subordinated bonds. While the first column

[22] In further investigation that we do not tabulate, we find that matrix prices become especially problematic when bonds are in unusual situations, such as a deal nearing its anticipated maturity, or a bond's recovering from an interest shortfall.

shows that our estimates are statistically reasonable, the subsequent columns show that we predict the yields in this subsample less accurately although we still capture a substantial amount of variation. While in the second column, the correlation coefficient is .5505, the combination of the ratio of standard deviations (.9278) and the R^2_{OOS} (.1607) indicate that in this subsample our predictions are slightly noisier and in any case less accurate, than in the other subsamples. This is economically intuitive as variation in yields in this segment is higher, and transactions are more scarce and more heterogeneous. The third column shows that the mixed yield the period before the transaction does slightly worse here, with an R^2_{OOS} of .1246. This also indicates that the argument about matched yields described for senior bonds does not apply in this segment, which, again is consistent with intuition. Lastly, column 4 indicates that in this segment the noise problem with matrix prices is especially severe, with a ratio of standard deviations of nearly 2 and an R^2_{OOS} of -3.12. This highlights how difficult it is to price bonds in this market segment. In comparison, our estimation is still able to produce reasonable yields, even in this challenging environment.

Panel B shows a subsample of results post the financial crisis. Overall, our results are qualitatively maintained, with a slight increase in prediction accuracy from our mixed yields (column 3), for all bonds (R^2_{OOS} of .18 instead of .15) and a slight decrease in predictive accuracy for pure prediction (column 2). For senior bonds, the accuracy of the mixed yields (column 3 again) decreases and is now lower than that of pure predicted yields, which is largely unaffected. The predictions for subordinated bonds in this time period are of similar quality, with only slightly more noise for pure predictions (higher ratio of standard deviations, lower R^2_{OOS} and slightly less noise for matched yields). The noise problem with matrix prices is more severe in this subsample with higher ratios of standard deviations and more negative R^2_{OOS}. Overall, Table 5 shows that our *Self-Propagating Rolling Panel VAR* estimation procedure provides a quantitatively satisfactory model of the price formation process for the CMBS market, in an unrestricted setting.

4.3 Individual CMBS bonds: Restricted VAR

Table 6 shows results from estimating the rolling VAR with the Campbell-Shiller restrictions, which allow only cash flow information to drive yields. We report the same statistics as in the previous tables, omitting the column for Matrix vs Transaction yields, as these values do not change in this setting. In addition to the previous statistics, we also report DM, the value and corresponding significance level of a Diebold-Mariano test with Harvey, Leybourne, Newbold modification of the null hypothesis that the prediction errors from the restricted VAR equal those from the cash flow only specification of the unrestricted VAR (whose results are reported in Table 5).

For all bonds in the full sample (Panel A, top part), we see that all statistics reported closely resemble those for the unrestricted VAR, with reasonable statistical properties when compared to matrix prices. Similarly, the ratios of σ for *Pred vs. Trans* and *Mixed vs. Trans* are close, but actually a few percentage points higher (indicating a slightly noisier model, when seen in conjunction with the other statistics). The correlation point estimates are actually slightly *higher* than for the unrestricted model, while the R^2_{OOS} statistics are slightly lower, with the largest drop seen in the *Pred vs. Trans* values. The DM statistics in both cases, however, indicate that the prediction errors from the two models are not statistically distinguishable. This provides strong evidence in favor of the hypothesis that cash flow information constitutes an important driver of yields in this market: the unrestricted and restricted VAR have the same predictive power, meaning that the pricing process in the market naturally follows the Campbell-Shiller restrictions.

This picture changes as we subdivide our sample and conduct separate VAR runs for senior (AAA) and subordinated (non-AAA) bonds. For the senior bonds in Panel A, we find a drop in predictive power for *Pred vs. Trans* with R^2_{OOS}, dropping from .2254 to .1781, and the DM statistic indicating that this is strongly statistically significant. For *Mixed vs. Trans* the drop in predictive power is much smaller, with R^2_{OOS} going from .2322 to .2317, and the DM statistic only rejecting at the 10% level.

These results are economically intuitive. The Campbell-Shiller restriction forces the VAR to generate yields that are driven only by cash flow information. In the CMBS market, even through and after the financial crisis, there were very few shortfalls in senior bonds. This is due to their credit protection: nearly all shortfalls driven by mortgage delinquencies were absorbed by the subordinate bonds. In other words, the subordination structure functioned as it was supposed to. What did occur, as subordinate tranches experienced shortfalls, however, was a yield expansion in senior bonds, to reflect downgrade risk and reduced distance to default.

Figure 1 and Table 7 illustrate this point. Note that throughout the time of the financial crisis (2008 and part of 2009) delinquencies rise steadily. From Table 7 we see that yields follow suit, for both senior and subordinated bonds (though much more so for the latter). Notice, however, that cash flow shortfalls throughout the crisis remain extremely minor for both senior and subordinated bonds (only .06% of senior bonds and .43% of subordinated bonds showed any shortfall whatsoever). For subordinated bonds, shortfalls rose dramatically around- and after 2010, while for senior bonds there never were large amounts of shortfalls at all.

When estimating the unrestricted Campbell-Shiller VAR, one also allows past yields as an additional driver of yields. For this reason, the unrestricted VAR has significantly more explanatory power than the restricted VAR, since yields expanded in a persistent way. Senior cash flows, on the other hand, remained nearly fully intact and so are unable to predict this yield expansion. The fact that the drop in predictive power is far less pronounced in *mixed* yields is due to the fact that the construction of mixed yields use transaction-based yields of matched bonds, and only use VAR predictions when no transactions are available. The matched yields, of course, track the yield expansion in senior yields as they are not limited by a restriction that limits these to be cash flow driven, and therefore capture a much larger portion of yield variation.

The next section of Table 6 shows a similar picture for non-AAA (i.e. subordinate) bonds. Here, for *Pred vs. Trans*, the R^2_{OOS} drops dramatically from .1607 to nearly zero (.0072), and once again the DM statistic shows this drop to be significant. This result, at first, seems puzzling, since subordinate bonds did suffer cash flow shortfalls in the financial crisis. However, as discussed above, even in subordinate bonds, mortgage delinquencies took a while to work themselves through the system. This means that subordinate bondholders were still not experiencing significant shortfalls, even after the financial crisis set in. During this entire time, however, bondholders were aware of the onset of delinquencies (either in their own deal, or in comparable deals which they observed, or in the market as a whole), and adjusted yields for the risk that they, too, would be affected by delinquencies. The restricted VAR, once again, is unable to forecast this since at the beginning of the financial crisis yields expanded without being preceded by a cash flow shortfall. These would be necessary for the restricted VAR to predict yield expansions. The unrestricted VAR, on the other hand, uses the persistence of the yield expansion as additional information to forecast yields and therefore has significantly more predictive power. Similar to the situation for the AAA bonds, for the *mixed* yields, the matched yields that are used in the panel are full yields (and not only cash flow based) and thus capture this yield expansion.[23]

The fact that, despite the effects described, the VAR does not lose significant predictive power over the entire set of bonds (top section of Panel A) suggests that information, and especially cash flow related information, from subordinate bonds helps price senior bonds. The likely mechanism at work is that senior bondholders can observe cash flow shortfalls in subordinated bonds within their own deal (or in deals that are similar to their own), and use this information to expand yields, to reflect downgrade risk or reduced distance to default in their own bonds. [24] Since there are many more senior bonds than subordinate bonds in the sample, and all these are priced better when using the information from subordinate bonds in conjunction, the pricing power over the entire cross section is not significantly affected by the restriction.

Comparing Panel B (which examines only the post-crisis period), with Panel B of Table 5, we find no drop in predictive power between the unrestricted- and restricted VAR for all bonds, with R^2_{OOS} values actually a few points *higher* for the restricted VAR, and the DM statistic failing to reject.[25] For senior bonds, we find a very similar picture to the one we see over the entire time period with a larger and significant drop in R^2_{OOS} for *Pred vs. Trans* and a smaller one, only significant at the 10% level for *Mixed vs. Trans.* The effect described above still applies here, in that even throughout- and after the financial crisis senior CMBS bonds never suffered substantial cash flow shortfalls that could predict yields in a restricted setting.

[23]While we motivate results obtained over the entire time sample with effects caused before- and during the crisis, this should still be plausible. Recall that, even though our first predicted yields are for 2006, we *train* the rolling VAR over a three-year window starting in 2003. Since cash flow shortfalls did not start in significant amounts until about 2009, we would not see these used for predicting yields until a large enough part of the VAR's estimation window contains them. This creates a reduced predictive power over large-enough portions of our time period to produce the overall results we have. The fact that we are in a panel with a large cross section does speed up this incorporation process, but a definite lag should still be visible.

[24]We explicitly document this effect in Section 4.5.

[25]The higher point estimate could be due to estimation noise in this case; since this is a rolling out-of-sample estimation, there should be no mechanical effect that necessarily makes point estimates for the accuracy of yield prediction lower.

For the subordinate bonds in the bottom section of the table, on the other hand, the picture does change after the crisis. Now we see predictive power by the restricted VAR that is not statistically different from that generated by the unrestricted VAR. In this period there were cash flow shortfalls throughout this class of bonds, and these were used by the market in forming yields. The fact that both here, as well as for all bonds, the restricted VAR has the same predictive power as the unrestricted VAR constitutes strong evidence that the market naturally follows the Campbell-Shiller restrictions in forming yields, which means that cash flow information constitutes an important driver of yields. However, very plausibly, this mechanism only works in times when there is variation in cash flows.

4.4 A Model with Added Discount Rate Variables

To complete our picture, we now investigate further the role of discount factor information in this setting. To do so, we investigate the added predictive power of the unrestricted *self-propagating* VAR when specific discount factor variables are added to this model. The two discount-factor variables we find most useful here are squared cash flow growth (which captures instantaneous volatility) and REIT returns (which captures a time-varying real-estate specific discount factor in an easily observable way).[26]

Table 8 presents these results. As before, Panel A shows the results for the full time series, while Panel B shows results for the post-crisis period. Within each panel, once again we show results for all bonds, senior bonds (AAA at origination) and subordinate bonds (Non-AAA at origination). Comparing these results to those in Table 5 shows a substantial increase in predictive power for the full set of bonds, especially for pure predictions (i.e. $Pred_t vs. Trans_t$). The R^2_{OOS} here is .2881, compared to the .1964 for the cash flow only specification. At the same time, the ratio of standard deviations is actually *lower* here (.7734 versus .9112), indicating less estimation noise. For $Mixed_{t-1} vs. Trans_t$, we also find a somewhat increased explanatory power with an R^2_{OOS} of

[26] In untabulated results, we try a variety of other discount factor proxies but find that these largely add noise to our out-of-sample predictions, without improving explanatory power. The variables we have tried include (in no particular order): market-wide AAA and Baa credit spreads (from the Federal Reserve); various measurements for level of delinquency within a deal (from Trepp); covariance of market-wide CMBS cash flows with the S&P 500 (to proxy for Beta), aggregated at both a value-weighted and equal-weighted basis.

.1948 here versus .1515 for the cash flow only specification. Similarly, there is a slight drop in ratio of standard deviations (.89 here, versus .96 in Table 5).

Unlike the results for all bonds, the results for the set of senior and subordinated bonds in Panel A are very similar to those in the cash flow only specification, with variations of only a percentage point or less in the R^2_{OOS} values and similarly small variations in the ratios of standard deviations. Panel B shows a similar picture in post-crisis results, with substantial increases in predictive power for the set of all bonds (still compared to Table 5), and largely unchanged results for each subset.

The picture that emerges thus, is that discount factor variables help explain yields, but only over the set of all bonds, and not within each category. The intuition behind this result is, once again, that an information flow exists between subordinate bond cash flows and senior bond discount factors. The additional discount factor variables could act to help stabilize coefficients and therefore reduce noise, in such a way as to allow this information to be incorporated in yield predictions for senior bonds when estimating the cross section jointly. Without this information flow, on the other hand, the discount factor variables do not seem to have much of an effect.

4.5 Information Flow from Subordinate to Senior Bonds

We now investigate this phenomenon more directly. The economics are as follows: subordinate bonds in a deal provide credit protection to senior bonds. However, the mortgage pool underlying both types of bonds in the same deal is identical. Therefore, if subordinate bonds suffer cash flow shortfalls, this should serve as a warning to senior bond holders, of increased risk of shortfalls in their own bonds. This understanding should then lead to an increase in senior bond yields to account for this risk. This information transmission mechanism should also happen across deals that contain similar mortgages, in that cash flow shortfalls in a mortgage pool different from one's own, but with similar make-up may provide information about impending cash flow shortfalls in one's own pool.

To investigate this empirically, we form value-weighted portfolios of senior- and subordinated bonds, by deal vintage.[27] We then regress the log of value-weighted average yield from each portfolio

[27] Since CMBS deals contain loans of various sizes, property types, and geographic areas, vintage is the common way in which market participants categorize deals as *similar*.

of senior bonds of vintage v at time t, on the log of cash flow level at the same time and up to four lags of log cash flow growth from subordinate bonds of the same vintage, in the following panel regression:

$$\begin{aligned} senior.yield_{t,v} &= \alpha + \gamma_1 subord.cf_{t,v} + \gamma_2 \Delta subord.cf_{t,v} + \gamma_3 \Delta subord.cf_{t-1,v} + \\ & \quad \gamma_4 \Delta subord.cf_{t-2,v} + \gamma_5 \Delta subord.cf_{t-3,v} + \gamma_6 \Delta subord.cf_{t-4,v} \end{aligned} \tag{20}$$

If there is an information flow from subordinate cash flows to senior yields, we expect to find a positive coefficient for cash flow level ($subord.cf_{t,v}$) and negative coefficients for all cash flow growth variables. We include up to four lags to allow for the gradual propagation of information. We estimate this panel regression with- and without date fixed effects. All variables are in logs.

Table 9 shows the results from this regression. The cash flow level of subordinate bonds has a positive and significant coefficient in all specifications and over all time period subsamples. Cash flow growth variables have negative and significant coefficients. The absolute coefficient size (as well as the t statistics) decrease as the lag length increases, consistent with a gradual decay of information over time. Also of note are the high $\overline{R^2}$ statistics. These indicate that this model accounts for one third- to more than half of the total variation of senior yields. This supports the hypothesis that cash flow information from subordinate bonds plays an important role in pricing senior bonds.

5 Conclusion

The CMBS market differs from the stock market. While stocks tend to trade in liquid markets and receive dividends subject to firms' business prospects, CMBS bonds are thinly traded and receive contractually specified cash flows. Thus, the CMBS market is characterized by poor pricing information and good cash flow information. Previous research has found that discount rate information predominates the stock price formation process. To the contrary, we find that cash flow information drives the CMBS bond price process both within- and across bond classes.

We reach our conclusions by devising a *Self-Propagating* Rolling Window Panel VAR technique, based on a dynamic Gordon model, with which we predict yields of CMBS bonds in between trades. Using this methodology, we then find that while discount rate information helps explain some variation in CMBS bond yields, cash flow information plays a substantial role as a driver of volatility in this market, both at the aggregate- as well as at the individual-bond level.

Although we apply our methodology to CMBS bonds, it is applicable to any asset that is thinly traded but has consistent cash flow information. Other examples of such asset classes include commercial real estate, natural resource extraction sites (such as mines or oil and gas wells), as well as thinly traded fixed income securities with variable cash flows, such as municipal revenue-based bonds. Such asset classes offer opportunities for further research in this area.

References

An, X., Deng, Y. and Gabriel, S. A.: 2011, Asymmetric information, adverse selection, and the pricing of cmbs, *Journal of Financial Economics* **100**(2), 304–325.

Campbell, J. Y.: 1991, A variance decomposition for stock returns, *The Economic Journal* **101**(45), 157–179.

Campbell, J. Y. and Ammer, J.: 1993, What moves the stock and bond markets? a variance decomposition for long-term asset returns, *The Journal of Finance* **48**(1), 3–37.

Campbell, J. Y. and Shiller, R. J.: 1988a, The dividend price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies* **1**(3), 195–228.

Campbell, J. Y. and Shiller, R. J.: 1988b, Stock prices, earnings and expected dividends, *Journal of Finance* **43**(3), 661–676.

Childs, P. D., Ott, S. H. and Riddiough, T. J.: 1996, The pricing of multiclass commercial mortgage-backed securities, *Journal of Financial and Quantitative Analysis* **31**(4), 581–603.

Diebold, F. X. and Mariano, R. S.: 1995, Comparing predictive accuracy, *Journal of Business & Economic Statistics* **13**(3).

Greene, W. H.: 2003, *Econometric Analysis*, Upper Saddle River, NJ: Pearson Education, Inc.

Harvey, D., Leybourne, S. and Newbold, P.: 1997, Testing the equality of prediction mean squared errors, *International Journal of forecasting* **13**(2), 281–291.

Henningsen, A. and Hamann, J. D.: 2007, Systemfit: A package for estimating systems of simultaneous equations in r, *Journal of Statistical Software* **23**(4), 1–40.

Lochstoer, L. A. and Tetlock, P.: 2017, What drives anomaly returns? Working Paper.

Love, I. and Zicchino, L.: 2006, Financial development and dynamic investment behavior: Evidence from panel var, *The Quarterly Review of Economics and Finance* **46**(2), 190–210.

Mühlhofer, T. and Ukhov, A. D.: 2011, Do stock prices move too much to be justified by changes in dividends? evidence from real estate investment trusts. Working Paper, Available at http://tobias.muhlhofer.com.

Petersen, M. A.: 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of financial studies* **22**(1), 435–480.

Shiller, R. J.: 1992, *Market Volatility*, The MIT Press.

Titman, S. and Tsyplakov, S.: 2010, Originator performance, cmbs structures, and the risk of commercial mortgages, *The Review of Financial Studies* **23**(9), 3558–3594.

Vuolteenaho, T.: 2002, What drives firm-level stock returns?, *The Journal of Finance* **57**(1), 233–264.

Welch, I. and Goyal, A.: 2008, A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *The Review of Financial Studies* **21**(4), 1455–1508.

Wooldridge, J. M.: 2010, *Econometric analysis of cross section and panel data*, The MIT press.

Table 1: Summary Statistics for CMBS data

This table presents summary statistics for the CMBS data. Each panel shows the number of bonds per segment of initial rating. Then, Panel A shows balances, Panel B shows subordination levels, Panel C shows coupon rates, and Panel D shows weighted average life.

Panel A: Balance

Rating	**Number of Bonds**	**Cumulative Balance**	**Mean**	**Std**	**Min**	**Max**
AAA	2064	593,198,056,562	287,402,159	287,310,543	12,500,000	3,661,032,000
$AA+$ to $BBB-$	1779	56,310,875,243	31,653,105	21,735,164	1,050,000	274,733,000
$< BBB-$	138	2,999,639,115	21,736,515	20,580,732	1,819,000	142,700,000
Total	3981	652,508,570,920				

Panel B: Subordination (Percent of Total Balance)

Rating	**Number of Bonds**	**Mean**	**Std**	**Min**	**Max**
AAA	2064	23.24	6.82	5.625	32
$AA+$ to $BBB-$	1779	9.83	4.12	2.25	26.25
$< BBB-$	138	4.6	1.95	0	10

Panel C: Coupon (Percent of Total Balance)

Rating	**Number of Bonds**	**Mean**	**Std**	**Min**	**Max**
AAA	2064	4.97	1.15	0.662	7.95
$AA+$ to $BBB-$	1779	5.68	0.86	2.57	8.55
$< BBB-$	138	5.71	0.93	2.972	7.85

Panel D: Weighted Average Life (Years)

Rating	**Number of Bonds**	**Mean**	**Std**	**Min**	**Max**
AAA	2064	7.57	2.34	1.98	12.2
$AA+$ to $BBB-$	1779	10.13	0.96	4.38	17.4
$< BBB-$	138	11.14	2.01	9.13	17.57

Table 2: Results from Decompositions of Single-Time Series VAR Predictions

This table shows correlations between the component of predicted long-term yield attributable to cash-flow growth according to the Campbell-Shiller decomposition (δ_{ct}), and actual realized yields. We show correlations for all bonds, bonds rated AAA at origination, and bonds not rated AAA at origination, in each case re-estimating the VAR for the subset of bonds. We show first correlations over the full time period. Then we show these correlations over a *Pre-Crisis* period (up to the end of 2007) and a *Post-Crisis* period (from the beginning of 2009), to show changes before and after the financial crisis. For the time-period sub-samples, we first show correlations for δ_{ct} values computed by a single VAR run over the entire period, with only correlations computed over the sub-periods. Then we show correlations for δ_{ct} values generated by computing two separate VARs, one over each time-period subsample. Note that the full-sample VAR run and correlations contains the period of the financial crisis, while the two sub-samples do not. Significance stars are for the test of the null hypothesis of zero correlation.

δ_{ct}

Time Period	**All Bonds**	**AAA at Orig.**	**Non-AAA at Orig.**
Full Sample			
Full Time Period	0.82***	0.6481***	0.8999***
Time-Period Sub-Samples without Re-Estimating VAR			
Pre-Crisis	0.5476***	0.3938**	0.3504***
Post-Crisis	0.9256***	0.8092***	0.8532***
Time-Period Sub-Samples Re-Estimating VAR			
Pre-Crisis	0.3965**	0.346**	0.4377***
Post-Crisis	0.9926***	0.7958***	0.9913***

° : significance level < 10%. *: significance level < 5%. **: significance level < 1%. *** : significance level < 0.1%.

Table 3: Full-Sample Panel VAR Coefficients

This table shows VAR coefficients for our full-sample panel VAR. The state variables used are *yield*, the log of the cash-flow yield implied by matrix prices, *lt.rate*, the log of the long-term interest rate, and *cf.growth*, the change in monthly log of bond cash flow. The table shows the coefficients for the VAR using one lag.While we run these VARs for up to four lags, we omit tabulating the coefficients for our higher-lag VARs, to save space. The VARs are estimated through equation-by-equation OLS.

Dependent	$yield_{t-1}$	$lt.rate_{t-1}$	$cf.growth_{t-1}$	$\overline{R^2}$	F
$yield_t$	0.9546	0.0153	−0.4299	0.8738	612592
	(1355.46)***	(6.75)***	(−222.19)***		
$lt.rate_t$	0.0064	0.9415	−0.0044	0.8936	742909
	(32.62)***	(1492.07)***	(−8.19)***		
$cf.growth_t$	−0.0376	0.0099	−0.4357	0.2175	24598
	(−59.76)***	(4.92)***	(−252.01)***		

° : significance level < 10%. *: significance level < 5%. **: significance level < 1%. *** : significance level < 0.1%.

Table 4: Predictive Power of Panel Vector Autoregressions.

This table presents statistics comparing the predicted yields from our VARs, with actual ex-post realized yields. Specifically, the table presents, for each VAR specification, the ratio of the standard deviations between the predicted and the realized series of yields. Both sections present this for the predictions (i.e. fitted values) from a full-sample panel VAR. The variable $pred.yield_{i,t}$, is the fitted value from the VAR for this variable. The realized (matrix-price based) yield is $yield_{i,t}$. The set of state variables for the VAR consists of $yield_{i,t}$, $lt.rate_t$, the log of the long-term interest rate, and $cf.growth_{i,t}$, the monthly difference of bond cash flows. In addition, the second system contains $cf.growth^2_{i,t}$, the log of the squared difference of bond cash flow, and $reit.ret_t$, the return to the REIT market. All variables are de-meaned.

Lags:	**1**	**2**	**3**	**4**
Full-Sample Panel VAR System: $[yield_{i,t}, lt.rate_t, cf.growth_{i,t}]'$				
$\sigma(pred.yield_{i,t})/\sigma(yield_{i,t})$	0.9348	0.9417	0.9461	0.9483
Full-Sample Panel VAR System: $[yield_{i,t}, lt.rate_t, cf.growth_{i,t}, cf.growth^2_{i,t}, reit.ret_t]'$				
$\sigma(pred.yield_{i,t})/\sigma(yield_{i,t})$	0.9359	0.9432	0.9476	0.9496

Table 5: Predictive Power of Self-Propagating Transaction-Based Panel Vector Autoregressions, for CMBS Bonds, Cash-Flow-Only Specification.

This table presents statistics comparing the predicted CMBS-bond yields from our Self-Propagating Transaction-Based Rolling Panel VARs, with ex-post realized yields. The table first presents statistics that compare our predicted yields with matrix-price-based yields. Then, the table shows how our predicted yields compare with ex-post realized transaction-based yields in the period of each transaction. We then show how our *Mixed* yields one month before the transaction compare to transaction yields, and for benchmarking, how the matrix-price yield the period before the transaction compares to the transaction yield. The statistics we present are ratios of standard deviations of the predicted over the realized series, correlation coefficients between each pair of series, with t-statistics for the null hypothesis that the actual correlation between the two series is 0 in parentheses, and lastly, for comparison with actual transaction-based yields, an out-of-sample R^2 (R^2_{OOS}). The set of state variables for this VAR consists of only $mixed.yield_{i,t}$, $lt.rate_t$, the log of the long-term interest rate, and $cf.growth_{i,t}$. Mixed yield consists of logs of transaction yields (where they occur), matched yields, where transactions occur for a bond of the same vintage and initial rating, and predicted yield from the rolling VAR system where they do not. The window size for the rolling VAR is 36 months. Panel A presents results from the entire available time period (2006-2013), while the subsequent panel shows a later subset, as labeled. The frequency is monthly.

Panel A: Full Sample

Measure	**Pred$_t$ vs. Mat$_t$**	**Pred$_t$ vs. Trans$_t$**	**Mixed$_{t-1}$ vs. Trans$_t$**	**Mat$_{t-1}$ vs. Trans$_t$**
		All Bonds		
Ratio of σ	0.3406	0.9112	0.9599	1.8882
Correlation	0.1788	0.5634	0.5548	0.246
t-statistic	(80.54)***	(99.69)***	(96.62)***	(36.77)***
R^2_{OOS}		0.1964	0.1515	−2.5804
		AAA at Origination		
Ratio of σ	0.3955	0.4538	0.543	1.2316
Correlation	0.4834	0.4753	0.48	0.3832
t-statistic	(174.22)***	(73.07)***	(73.35)***	(55.61)***
R^2_{OOS}		0.2254	0.2322	−0.5508
		Non-AAA at Origination		
Ratio of σ	0.3335	0.9278	0.9589	1.9992
Correlation	0.1457	0.5505	0.5401	0.2027
t-statistic	(44.53)***	(35.4)***	(34.13)***	(11.01)***
R^2_{OOS}		0.1607	0.1246	−3.1212

$^{\circ}$: significance level $<$ 10%. *: significance level $<$ 5%. **: significance level $<$ 1%. *** : significance level $<$ 0.1%.

Panel B: Post-crisis (2009 forward)

Measure	**Pred**$_t$ **vs. Mat**$_t$	**Pred**$_t$ **vs. Trans**$_t$	**Mixed**$_{t-1}$ **vs. Trans**$_t$	**Mat**$_{t-1}$ **vs. Trans**$_t$
		All Bonds		
Ratio of σ	0.3285	0.9781	0.9345	2.0648
Correlation	0.1671	0.5679	0.5287	0.2086
t-statistic	(64.49)***	(88.4)***	(79.11)***	(27.1)***
R^2_{OOS}		0.1539	0.1776	−3.2738
		AAA at Origination		
Ratio of σ	0.3687	0.4347	0.5153	1.4085
Correlation	0.4835	0.4625	0.3927	0.3172
t-statistic	(154.07)***	(62.21)***	(50.47)***	(39.53)***
R^2_{OOS}		0.2129	0.1767	−1.0388
		Non-AAA at Origination		
Ratio of σ	0.3172	0.9221	0.8872	2.1583
Correlation	0.1296	0.5453	0.4885	0.1669
t-statistic	(33.57)***	(30.27)***	(25.87)***	(7.82)***
R^2_{OOS}		0.1547	0.1488	−3.8132

° : significance level < 10%. *: significance level < 5%. **: significance level < 1%. *** : significance level < 0.1%.

Table 6: Predictive Power of Self-Propagating Transaction-Based Panel Vector Autoregressions, for CMBS Bonds, Cash-Flow-Only Specification, Restricted VAR.

This table presents statistics comparing the predicted CMBS-bond yields from our Self-Propagating Transaction-Based Rolling Panel VARs, with ex-post realized yields. In this table, the VAR coefficient matrix is estimated according to the Campbell-Shiller cross-equation restrictions. The table first presents statistics that compare our predicted yields with matrix-price-based yields. Then, the table shows how our predicted yields compare with ex-post realized transaction-based yields in the period of each transaction. We then show how our *Mixed* yields one month before the transaction compare to transaction yields. The statistics we present are ratios of standard deviations of the predicted over the realized series, correlation coefficients between each pair of series, with t-statistics for the null hypothesis that the actual correlation between the two series is 0 in parentheses, and lastly, for comparison with actual transaction-based yields, an out-of-sample R^2 (R^2_{OOS}). DM is a Diebold-Mariano test statistic (with the Harvey, Leybourne, and Newbold modification), for the null hypothesis that the prediction errors from the restricted VAR and the unrestricted VAR are identical, against the two-sided alternative. The set of state variables for this VAR consists of only $mixed.yield_{i,t}$, $lt.rate_t$, the log of the long-term interest rate, and $cf.growth_{i,t}$. Mixed yield consists of logs of transaction yields (where they occur), matched yields, where transactions occur for a bond of the same vintage and initial rating, and predicted yield from the rolling VAR system where they do not. The window size for the rolling VAR is 36 months. Panel A presents results from the entire available time period (2006-2013), while the subsequent panel shows a later subset, as labeled. The frequency is monthly.

Panel A: Full Sample

Measure	**Pred$_t$ vs. Mat$_t$**	**Pred$_t$ vs. Trans$_t$**	**Mixed$_{t-1}$ vs. Trans$_t$**
		All Bonds	
Ratio of σ	0.3678	1.003	0.9785
Correlation	0.1777	0.5853	0.5721
t-statistic	(80)***	(105.52)***	(101.06)***
R^2_{OOS}		0.1657	0.1699
DM		0.5659	-0.4682
		AAA at Origination	
Ratio of σ	0.3956	0.5504	0.5437
Correlation	0.4832	0.4437	0.4797
t-statistic	(174.13)***	(66.97)***	(73.28)***
R^2_{OOS}		0.1781	0.2317
DM		7.3585***	1.661°
		Non-AAA at Origination	
Ratio of σ	0.332	1.0606	0.9398
Correlation	0.1504	0.5361	0.5524
t-statistic	(46.01)***	(34.09)***	(35.24)***
R^2_{OOS}		0.0072	0.1615
DM		2.0548*	-1.1417

$^{\circ}$: significance level $< 10\%$. *: significance level $< 5\%$. **: significance level $< 1\%$. *** : significance level $< 0.1\%$.

Panel B: Post-crisis (2009 forward)

Measure	**Pred$_t$ vs. Mat$_t$**	**Pred$_t$ vs. Trans$_t$**	**Mixed$_{t-1}$ vs. Trans$_t$**
		All Bonds	
Ratio of σ	0.3539	1.0067	0.9489
Correlation	0.1643	0.5856	0.5401
t-statistic	(63.37)***	(92.57)***	(81.52)***
R^2_{OOS}		0.1627	0.1849
DM		-0.1376	-0.1513
		AAA at Origination	
Ratio of σ	0.3688	0.4785	0.5164
Correlation	0.4832	0.4351	0.3923
t-statistic	(153.97)***	(57.63)***	(50.41)***
R^2_{OOS}		0.1804	0.176
DM		7.6123***	1.6771°
		Non-AAA at Origination	
Ratio of σ	0.3117	0.9104	0.8534
Correlation	0.1323	0.549	0.4941
t-statistic	(34.28)***	(30.57)***	(26.26)***
R^2_{OOS}		0.1638	0.1817
DM		-0.1629	-0.8543

$^{\circ}$: significance level $< 10\%$. *: significance level $< 5\%$. **: significance level $< 1\%$. *** : significance level $< 0.1\%$.

Table 7: Cash Flows and Yields around the Financial Crisis

This table shows differences in various bond measures between the end of 2007 (the start of the financial crisis) and the end of 2009. *Delinquency Rate* is the monthly (quarterly) average of the deal-level 60+ day delinquency rate (i.e. dollar amount delinquent, divided by total outstanding balance) for deals in the sample in a given month (quarter) (1.0=1%). *AAA Yield Difference* is the monthly (quarterly) average of the difference between the annualized traded yield on AAA bonds (actual coupon cash flow, divided by market price) and the annualized yield implied by their coupon rate (i.e. contractual coupon cash flow divided by face value outstanding) (1.0=1%). *non-AAA Yield Difference* is the monthly (quarterly) average of the difference between the annualized traded yield on non-AAA bonds and the annualized yield implied by their coupon rate (1.0 = 1%). *AAA % Interest Shortfall* is the percentage of AAA bonds in a given month (quarter) that are experiencing an interest shortfall (1.0=1%). *non-AAA % Interest Shortfall* is the percentage of non-AAA bonds in a given month (quarter) that are experiencing an interest shortfall (1.0 = 1%). *AAA* refers to bonds that were rated AAA at issuance, and are therefore the most senior bonds. *non-AAA* refers to subordinated bonds. Cash flow and delinquency data are from Trepp. Bond prices are from SNL Finanical and the NAIC.

Period	**Delinquency Rate**	**AAA Yield Difference**	**non-AAA Yield Difference**	**AAA % Interest Shortfall**	**non-AAA % Interest Shortfall**
Dec 2007	0.479	0.0749	0.2173	0.064	0.434
Dec 2009	1.8	1.661	25.34	0.064	0.374
Q4 2007	0.409	0.0956	0.2255	0.066	0.376
Q1 2009	2.083	1.4337	25.6462	0.065	0.313

Table 8: Predictive Power of Self-Propagating Transaction-Based Panel Vector Autoregressions, for CMBS Bonds, Full Specification.

This table presents statistics comparing the predicted CMBS-bond yields from our Self-Propagating Transaction-Based Rolling Panel VARs, with ex-post realized yields. The table first presents statistics that compare our predicted yields with matrix-price-based yields. Then, the table shows how our predicted yields compare with ex-post realized transaction-based yields in the period of each transaction. We then show how our *Mixed* yields one month before the transaction compare to transaction yields. The statistics we present are ratios of standard deviations of the predicted over the realized series, correlation coefficients between each pair of series, with t-statistics for the null hypothesis that the actual correlation between the two series is 0 in parentheses, and lastly, for comparison with actual transaction-based yields, an out-of-sample R^2 (R^2_{OOS}). The set of state variables for the VAR consists of $mixed.yield_{i,t}$, $lt.rate_t$, the log of the long-term interest rate, $cf.growth_{i,t}$, the difference in log of bond cash flows, $cf.growth^2_{i,t}$, the log of the squared difference of bond cash flow, and $reit.ret_t$, the return to the REIT market. Mixed yield consists of transaction yields (where they occur), matched yields, where transactions occur for a bond of the same vintage and initial rating, and predicted yield from the rolling VAR system where they do not. The window size for the rolling VAR is 36 months. Panel A presents results from the entire available time period (2006-2013), while the subsequent panel shows a later subset, as labeled. The frequency is monthly.

Panel A: Full Sample

Measure	**Pred$_t$ vs. Mat$_t$**	**Pred$_t$ vs. Trans$_t$**	**Mixed$_{t-1}$ vs. Trans$_t$**
		All Bonds	
Ratio of σ	0.372	0.7734	0.8878
Correlation	0.1536	0.5733	0.5502
t-statistic	(68.89)***	(102.29)***	(95.45)***
R^2_{OOS}		0.2881	0.1948
		AAA at Origination	
Ratio of σ	0.3954	0.4188	0.5429
Correlation	0.4834	0.4765	0.48
t-statistic	(174.23)***	(73.31)***	(73.34)***
R^2_{OOS}		0.2221	0.2321
		Non-AAA at Origination	
Ratio of σ	0.3323	0.9096	0.9661
Correlation	0.1464	0.5497	0.5361
t-statistic	(44.77)***	(35.33)***	(33.77)***
R^2_{OOS}		0.1724	0.1109

$^\circ$: significance level $< 10\%$. *: significance level $< 5\%$. **: significance level $< 1\%$. *** : significance level $< 0.1\%$.

Panel B: Post-crisis (2009 forward)

Measure	**Pred$_t$ vs. Mat$_t$**	**Pred$_t$ vs. Trans$_t$**	**Mixed$_{t-1}$ vs. Trans$_t$**
		All Bonds	
Ratio of σ	0.3074	0.8305	0.8466
Correlation	0.1654	0.5785	0.5254
t-statistic	(63.83)***	(90.87)***	(78.43)***
R^2_{OOS}		0.271	0.2397
		AAA at Origination	
Ratio of σ	0.3686	0.4048	0.515
Correlation	0.4835	0.4648	0.3926
t-statistic	(154.08)***	(62.59)***	(50.45)***
R^2_{OOS}		0.2113	0.1766
		Non-AAA at Origination	
Ratio of σ	0.3167	0.9217	0.9001
Correlation	0.1308	0.5436	0.486
t-statistic	(33.9)***	(30.14)***	(25.7)***
R^2_{OOS}		0.1525	0.1342

° : significance level $<$ 10%. *: significance level $<$ 5%. **: significance level $<$ 1%. *** : significance level $<$ 0.1%.

Table 9: Regressions of Senior Yields on Subordinate Cash Flows

Dependent variable: $senior.yield_{t,v}$. This table shows regressions of yields to senior bonds on cash-flow information to subordinate bonds. Specifically, for each vintage v and time t, for the dependent variable we compute the log of value-weighted average yields across all senior bonds backed by mortgages of that vintage. Similarly, for the independent variables, we compute logs of value-weighted average cash flows and growth rates of these, for all subordinate bonds of vintage v. We then regress yields on cash flow levels, and various lags of cash flow growth. The data is monthly. We present specifications with- and without date fixed effects. We show results for these regressions for the full time sample, as well as for the pre-crisis period (before 2006), the period of the financial crisis (2007-2009), and the post-crisis period (after 2009).

	Full Sample		**Pre-Crisis**		**Crisis**		**Post-Crisis**	
(Intercept)	0.0238	−0.4971	0.8162	0.7793	−0.1959	−0.0614	$-5e-04$	−1.5724
	(0.13)	(−2.14)*	(2.79)**	(2.54)*	(−0.65)	(−0.2)	(0)	(−4.2)***
$subord.cf_{t,v}$	1.018	0.9364	1.1773	1.1799	0.9654	1.0007	1.0126	0.6859
	(30.65)***	(22.33)***	(20.97)***	(20.51)***	(16.89)***	(17.61)***	(16.77)***	(9.78)***
$\Delta subord.cf_{t,v}$	−0.8802	−0.8063	−0.9473	−0.9503	−0.8114	−0.8485	−0.8894	−0.6018
	(−17.42)***	(−14.19)***	(−6.87)***	(−6.21)***	(−4.87)***	(−5.04)***	(−12.26)***	(−7.5)***
$\Delta subord.cf_{t-1,v}$	−0.7073	−0.6522	−0.657	−0.6625	−0.628	−0.6551	−0.7359	−0.5032
	(−12.12)***	(−10.24)***	(−3.77)***	(−3.48)***	(−3.15)**	(−3.24)**	(−9.41)***	(−5.94)***
$\Delta subord.cf_{t-2,v}$	−0.5574	−0.505	−0.5549	−0.5606	−0.4894	−0.5284	−0.5848	−0.3946
	(−9.22)***	(−7.82)***	(−2.99)**	(−2.7)**	(−2.5)*	(−2.66)**	(−7.43)***	(−4.73)***
$\Delta subord.cf_{t-3,v}$	−0.3955	−0.3476	−0.5157	−0.539	−0.2738	−0.3048	−0.4111	−0.2712
	(−6.89)***	(−5.73)***	(−2.54)*	(−2.37)*	(−1.69)°	(−1.83)°	(−5.65)***	(−3.56)***
$\Delta subord.cf_{t-4,v}$	−0.2004	−0.1672	−0.4989	−0.5338	−0.0832	−0.089	−0.2023	−0.1266
	(−4.28)***	(−3.41)***	(−2.26)*	(−2.14)*	(−0.68)	(−0.71)	(−3.53)***	(−2.14)*
$\overline{R^2}$	0.47	0.4878	0.6462	0.6317	0.4627	0.4964	0.312	0.3552
F	156.766	8.722	74.35	11.082	47.796	8.838	46.961	5.784
N	1055	1055	242	242	327	327	609	609
Date Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes

° : significance level < 10%. *: significance level < 5%. **: significance level < 1%. *** : significance level < 0.1%.

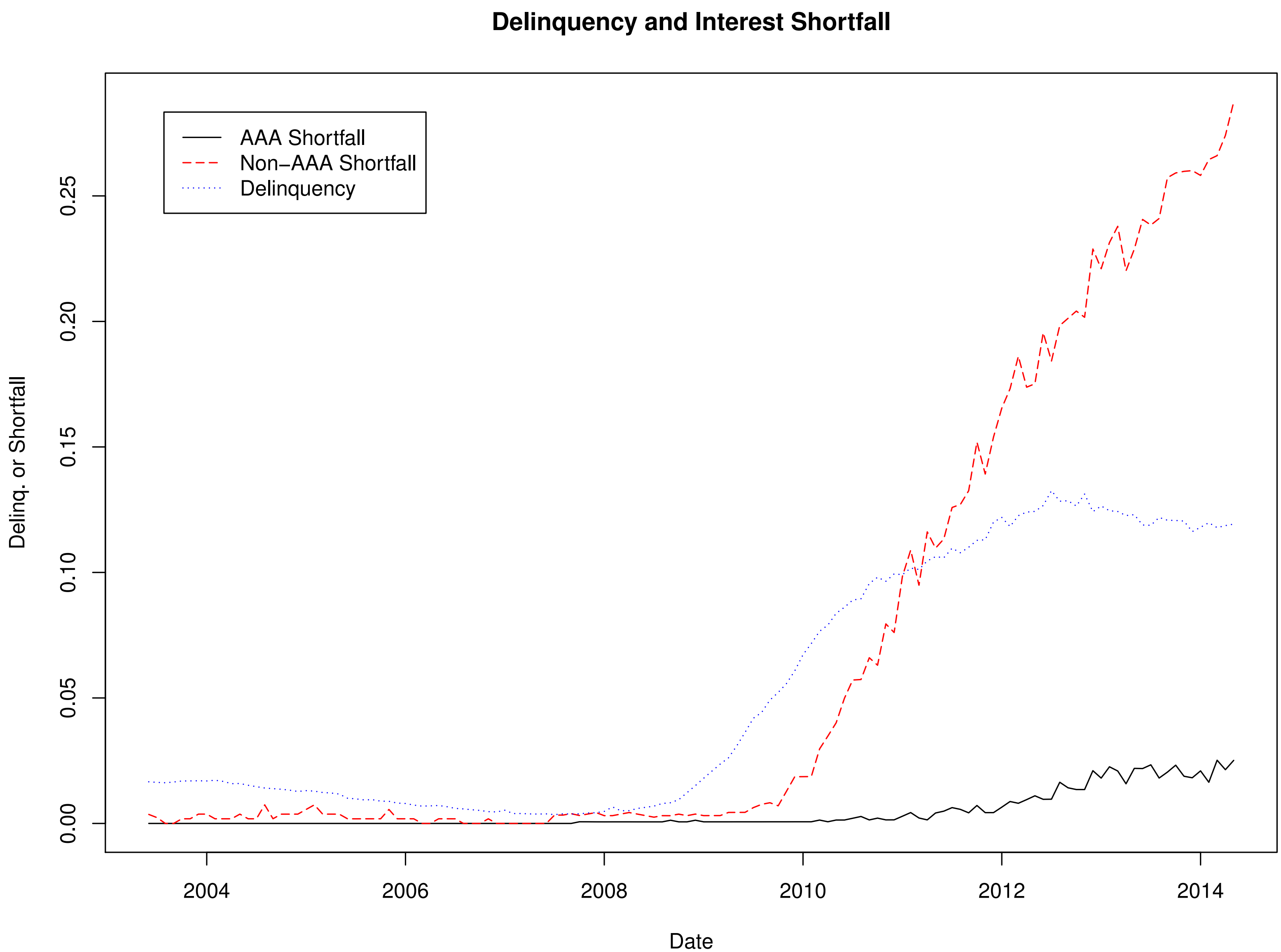

Figure 1: Delinquencies and Interest Shortfalls.

This figure shows time-series plots of Delinquency- as well as Shortfall rates. Specifically, *Delinquency* is the deal-level average quarterly delinquency rate (i.e. dollar amount which is 60+ days overdue, divided by total outstanding balance). *Shortfalls* is the fraction of bonds with any cash-flow shortfall. *AAA* refers to bonds rated AAA at inception (senior bonds) and *non-AAA* refers to subordinate bonds.