**Title**
The Market View of Mortgage Credit Risk



**Author**
O'Neill, James Edmund

**Publication Date**
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# The Market View of Mortgage Credit Risk

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Management

by

James Edmund O'Neill

2022

ABSTRACT OF THE DISSERTATION

The Market View of Mortgage Credit Risk

by

James Edmund O'Neill
Doctor of Philosophy in Management
University of California, Los Angeles, 2022
Professor Andrea Lynn Eisfeldt, Chair

In Chapter 1 of this dissertation, I study the informational content of GSE Credit Risk Transfer (CRT) bonds. CRT bonds amount to catastrophe bonds on underlying mortgage collateral, and are informative about the market price of credit risk for conforming mortgage loans. I analyze the information content of this new asset class, which holds effectively half of the default risk of the $12 trillion US mortgage market. I build a pricing model and extract default probabilities from a comprehensive hand-collected dataset on CRT bond issuances using TRACE pricing data. I find that while the guaranty-fee (g-fee) implied by early CRT issuances was high and near the level charged by the GSEs (30 bp), the market has stabilized in recent years and implied g-fees have fallen significantly to around 10-20 basis points. I discuss the implications of these findings for several facets of the mortgage market.

In Chapter 2, I build and estimate a top-down portfolio credit model matched to a sample of GSE Credit Risk Transfer bond (CRT) tranche market prices. Two state variables, modeled as exogenous Poisson intensities, represent two sources of default risk affecting mortgages, which intuitively map into a level of routine mortgage defaults and the risk of catastrophic losses. The market views the (risk-neutral) probability of a catastrophic event, on the order of the 2008 housing market crisis, as occurring roughly once every 25 years. The implied guaranty-fee (g-fee) is on the order of 15 basis points over the sample, and both risk factors contribute roughly equally to this spread. I also analyze the level of credit-protection offered by the current CRT design in a simulation study, and discuss the conditions under which "credit risk transfer transfers credit risk."

The dissertation of James Edmund O'Neill is approved.

Tyler Stewart Muir

Barney P. Hartman-Glaser

Francis A. Longstaff

Andrea Lynn Eisfeldt, Committee Chair

University of California, Los Angeles

2022

*I am grateful for the unwavering support of Kelly.*

*I also thank my parents Michael and Stasia.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am thankful to my committee members, Barney Hartman-Glaser, Francis Longstaff, Tyler Muir, and especially Andrea Eisfeldt, my committee chair. They provided valuable feedback and pushed me to think harder about my research questions. I also benefited with discussions with Lars Lochstoer and Stavros Panageas, and enjoyed teaching with David Wessels and Jinyuan Zhang. I also thank Brad Badertscher, Zhi Da, and Chris Jones who were my original inspiration to pursue this degree.

I am also thankful to all the other Anderson PhD students I had the pleasure of interacting with over the course of the last five years, with special thanks to Paul Huebner, Salil Gadgil, Ed Kim, Denis Mokanov, Paul Momtaz, Gabriel Cuevas Rodriguez, Zach Sauers, Clinton Tepper, and Danyu Zhang.

Thank you to UCLA Anderson current and former staff, including Veronica Avila, Kristin Christian, Craig Jessen, Monica Pena, and Da‘Janay Ray for their help navigating the program.

Lastly, I thank the UCLA Graduate Division, the UCLA Anderson Finance area, and the UCLA Fink Center for Finance for generous funding and financial support during my time completing this research.

# VITA

| | |
|---|---|
| 2011–2015 | B.B.A Finance, University of Notre Dame, Notre Dame, Indiana |
| 2011–2015 | Glynn Family Honors Program, University of Notre Dame, Notre Dame, Indiana |
| 2014 | Financial Analyst Intern, eBay. Inc., San Jose, California |
| 2015–2016 | Junior Trader, Belvedere Trading LLC, Chicago, Illinois |
| 2016–2017 | M.Sc. Finance, University of Southern California, Los Angeles, California |
| 2016–2017 | Merit Scholarship, USC Marshall School of Business, Los Angeles, California |
| 2021 | Economist Intern, Amazon Web Services, Seattle, Washington |
| 2017–2022 | Fink Center Fellow, UCLA Anderson School of Management, Los Angeles, California |
| 2017–2021 | Fellowship Recipient, UCLA Anderson School of Management, Los Angeles, California |
| 2017–2022 | Teaching Assistant, UCLA Anderson School of Management, Los Angeles, California |

# CHAPTER 1

# The Market View of Mortgage Credit Risk

## 1.1 Introduction

This paper provides an empirical analysis of a new asset class, Credit Risk Transfer bonds (CRTs), that back the credit of conforming loans purchased by the Government Sponsored Enterprises, Fannie Mae and Freddie Mac (the GSEs). The prices investors are willing to pay for these bonds are revealing about how the market views credit risk in diversified pools of mortgages. With a potential for GSE exit from conservatorship and the return of private capital to mortgage markets, a thorough analysis of the informational content of the CRT bond program is needed.[1] Since mortgage credit risk has previously been largely untradeable in capital markets, with the exception of infrequently traded private label CMOs before the Global Financial Crisis, these bonds are the first comprehensive avenue by which to study this important risk. In this paper, I build and estimate a model to extract the information that this asset class contains with regards to mortgage risk; I find the guaranty-fee (g-fee) implied by CRT tranche prices to be around 15 basis points on average, lower than that charged by the GSEs, and correlated with corporate credit spreads.

To put the importance of this marketplace in perspective, as of late 2021, the US mortgage market had a size of $12 trillion in outstanding household debt balance. Agency MBS, that is, securities backed by loans that meet the standards set forth

1. See Richardson, Van Nieuwerburgh, and White (2017) for a summary of policy discussions regarding GSE reform.

for purchasing and securitization by the GSEs, account for roughly two-thirds of that amount.[2] CRT programs at Fannie Mae and Freddie Mac back roughly 75% of the credit risk of these mortgages. And the principal form of risk transfer undertaken by the GSEs are in the form of CRT bonds as described above.[3] This back of the envelope calculation implies that roughly half of total credit risk related to the US mortgage market, not just that of conforming mortgages, are now held by investors in the form of CRT bonds. With so much discussion in the public sphere about taxpayer risk and the privatization of the GSEs, it is quite remarkable that so much credit risk has been quietly transferred to private investors over just the last nine years with little fanfare both in the popular press and academic literature. It also underscores the importance of an interest in the pricing and economics of the CRT programs in the academic finance literature.

While the exact mechanics underlying the CRT market can be convoluted due to its structure, the economics behind the CRT bond program are simple. Purchasers of CRT bonds place capital in a trust that backs a portion of the credit risk represented by a diversified pool of mortgages purchased and held (or securitized) by the GSEs. When losses are suffered on these mortgage loans, principal balances in the trust are written down for the investor, with the write-down amount transferred to the GSE to compensate them for the credit loss in their portfolio. As payment for providing this capital, CRT bondholders receive periodic coupon payments, and the return of their principal if losses are not suffered on the entirety of their notional position. Importantly, these bonds are traded in a secondary marketplace where investors can actively express their views about the risk in the housing market.

My model extracts the risk-neutral distribution of losses implied by the prices

2. SIFMA , Urban Institute State of Housing Finance, January 2022

3. The GSEs have experimented with other forms of risk transfer, such as Credit Reinsurance Risk Transfer (CIRT) and others. This paper will focus only on the bond issuance programs.

of CRT bond tranches, similar to how investors use implied volatility or correlation measures in other asset markets. Naturally, the expectation of these losses amortized over the life of the loan pools maps into a measure of the credit cost portion of the GSEs guaranty-fee, (g-fee) which is measured as expected losses plus a risk premium component. I find that my measure of this credit cost portion varies between 10 and 30 basis points and has leveled out significantly as the CRT program has matured to the lower end of this spectrum. This result has important implications for the GSEs and the future of their role in mortgage markets. The market implied g-fee can serve as a signal for benchmarking this fee, which plays an important role in the provision of credit subsidies, and the smooth functioning of capital markets through credit-guaranteed MBS and lowered mortgage rates for borrowers. Knowing where this market rate lies in relation to charged levels, for which this paper offers a novel methodology, gives the GSEs a tool through which to judge the competitiveness of their g-fees. Most importantly, the resulting risk-neutral probabilities include a risk premium component, which allows me to explicitly identify an implied g-fee.

CRT bond tranche attachment and detachment points, the levels of credit losses at which CRT bond investors began to experience principal write-downs, have varied over the 121 unique bond issuances in my comprehensive sample of all standard CRT bond issuances to date. The model views the present value of losses incorporated into tranche market prices as revealing about different pieces of the probability density function of this implied loss distribution; they identify the probability that losses meet or exceed tranche detachment points. I use these moments to fit the risk-neutral distribution, which I choose to model using a scaled beta distribution. This specification has the advantage of requiring only two parameters, which is well suited to matching CRT issuances that contain between 2 and 5 tranches. Its support is also bounded, which meets the requirement that losses on loans must fall between zero and an upper bound specified by the total credit coverage of the CRT deal.[4] This also allows the model to abstract slightly from which pieces of the capital structure are sold in CRT deals; the CRT deals reveal an underlying continuous risk-neutral loss distribution. Key to this modeling choice is the idea that since tranches are effectively credit derivatives of the same underling mortgage pool, in the absence of arbitrage their prices must be revealing about the same underlying distribution.

The principal contribution of this paper is to offer a parsimonious and quick-to-estimate model that can be applied to CRT bond issuances in order to derive a market view of mortgage credit risk. In my model, I use a top-down credit approach in which mortgage payments and prepayments are modeled at the aggregate pool level, with defaults specified exogenously. This approach has been used in the valuation of mortgage-backed securities, or other pools of default-sensitive instruments (Giesecke, Goldberg, and Ding (2011), Diener, Jarrow, and Protter (2012), Sirignano and Giesecke (2019), Longstaff and Rajan (2008), Chernov, Dunn, and Longstaff (2017), Fleckenstein and Longstaff (2022)). This approach requires exponentially less parameters and is transparent; it lets the market tell us it's view with limited assumptions, rather than through the convoluted lens of Monte Carlo simulations and option-adjusted spreads based on the modelers view under the physical probability measure. Top-down approximations perform especially well when there is substantial homogeneity in the underlying assets, which is precisely the case for GSE conforming mortgage loans serving as CRT collateral.

The mortgage market in the United States is unique in that, until the advent of the CRT programs started in the early 2010's, the majority of mortgage credit risk was held by enterprises that were effectively monoline insurers. This has led

4. I allow losses to exceed the total credit coverage of the CRT issuance by 25 basis points; because the probability of loss on the most senior CRT tranche is so low, this assumption has very little effect on the estimated g-fee. Section 1.5.4 discusses the possibility of losses exceeding the credit coverage of the CRT bonds.

to the strange situation in which despite a size on the order of the US corporate bond market, we know effectively nothing about market based pricing of mortgage default risk. Furthermore, due to the GSEs convoluted relationship with the federal government, there is no accepted benchmark for whether the g-fees being charged are fair, too high, or too low.[5]. This problem was further worsened by opaque and risky securitizations that blew up during the subprime mortgage crisis, which led to skepticism about the trading of mortgage credit risk. The GSEs play a delicate role in this market, where they must balance offering credit risk subsidies to advance public housing agendas, maintaining an orderly MBS marketplace, and charging a fair market price for their guaranty.

As alluded to above, key to the GSE mandate is the g-fee that the they charge mortgage originators to bear credit risk; of course, this fee is directly passed onto borrowers in the form of higher borrowing rates. The g-fee, however, need not be set at an efficient or actuarially fair level if the principal role of the GSEs is to promote a particular housing agenda. For example, to encourage home ownership, the FHFA could direct the GSEs to lower g-fees below what the market would charge to guaranty a similar loan. Market prices of CRT bonds allow us to construct a measure of what this market-based fee would be, and then directly compare it to what the GSEs are charging.

For context, figure 1.1 below shows the average effective g-fee charged by Fannie Mae over the period 2000 to 2021. In the wake of the financial crisis, in the GSEs elected to raise the fee on several occasions.[6] The fee here, as reported in the Fannie Mae 10k reports, includes administrative fees regarding to GSE operations and the

5. This problem frequently surfaces in public debate on whether the g-fees are proper or not. For example, in 2012, the FHFA solicited comments on standardizing g-fees and potentially charging a state-level fee. Since, 2009, the FHFA has issued a yearly study on charged g-fees at Fannie Mae and Freddie Mac, the most recent of which can be found here.

6. For example, a 25 basis point upfront adverse market charge in the wake of the financial crisis. A timeline of g-fee developments dcan be found here.

securitization of loans. This portion is generally assumed to be around 10 basis points (Palmer (2017)), meaning that the credit portion of the fee is around 30-35 basis points. Unfortunately, levels of g-fees are not broken down into their particular risk buckets, so it is difficult to ascertain the exact fee that is being charged on loans similar to those serving as CRT collateral; thankfully, CRT collateral is generally representative of the majority of GSE loans. The fee can also vary over time depending on the composition of loans (refinance originations, for example, can have different fees).

Expanding on the historical discussion of the g-fee from above, figure 1.2 shows its components. General expenses and the 10 basis point TCCA adjustment make up effectively 20 basis points of the fee, which are referred to as the non-credit costs (Figure 1.1 is net of this TCCA adjustment). Prices of CRT tranches do not tell us anything about this portion of the g-fee, but is it generally regarded as not being

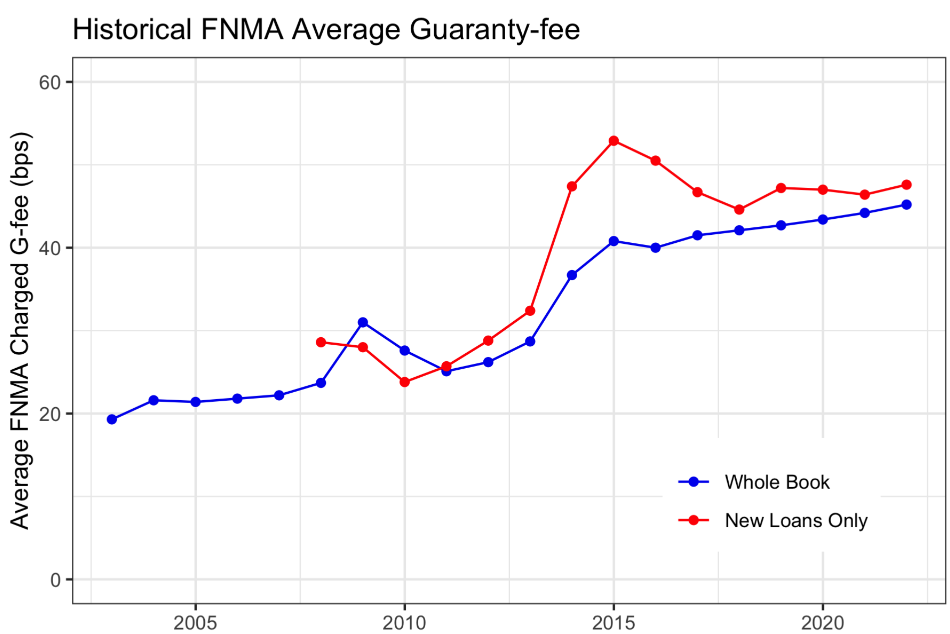



**Figure 1.1:** The g-fee charged by Fannie Mae has increased gradually since the financial crisis, and is now around 40-45 bps. The g-fee contains both an upfront and ongoing component. For the purposes of comparison, the upfront component is amortized over the typical life of a mortgage to generate a single value.

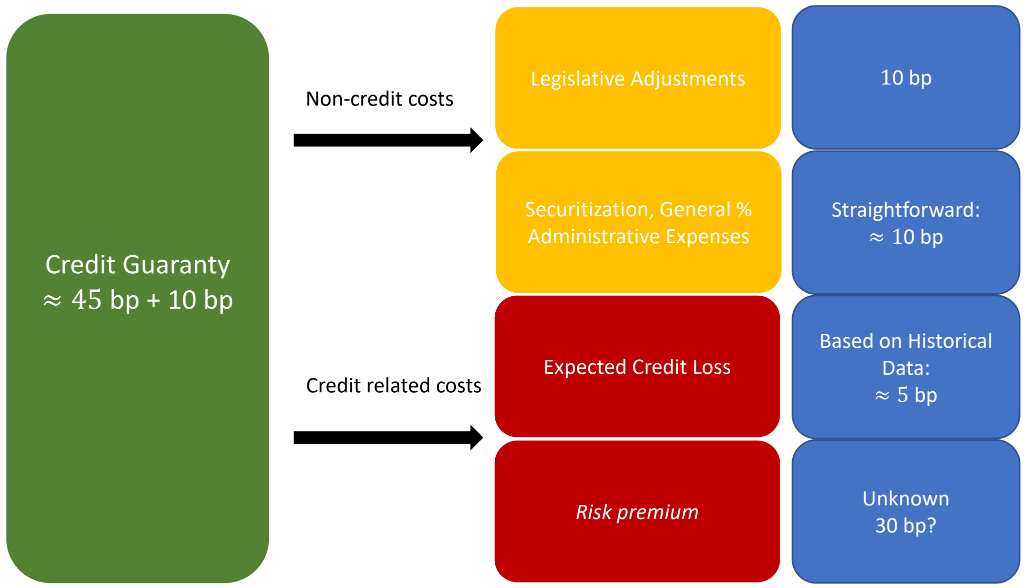



**Figure 1.2:** The g-fee can be broken down into costs associated with credit risk and costs that are not. The risk premium portion is particularly difficult to measure, especially from CRT bonds since they do not sell the catastrophic risk. For more information on the breakdown of the g-fee, see Palmer (2017).

subject to much debate. The second portion of the g-fee is referred to as the credit cost portion, which is an expected loss component plus a risk premium. Equation 2.12 breaks down this formula, where $CL$ represent annualized credit losses:

$$\text{g-fee} = SG\&A + \mathbb{E}^{\mathbb{P}}[CL] + RiskPremium = SG\&A + \underbrace{\mathbb{E}^{\mathbb{Q}}[CL]}_{\substack{\text{Identified by} \\ \text{CRT prices}}} \tag{1.1}$$

The risk-neutral losses implied by the CRT market can answer a variety of important economic questions about the housing market and role the GSEs play in it. First and foremost, as shown in the equation above the distribution of implied losses on a diversified mortgage pool can be directly mapped to the credit-cost portion of the fair g-fee that should be charged were the market offering to insure it, the role that the GSEs currently play in the US mortgage market. Unlike more naive methods, the methodology employed in this paper for estimating fair g-fees can speak to their distribution rather than simply their point estimate (Palmer (2017)). These results could be of use to the GSEs and regulators who want to assess the risk of mortgage portfolios in a value-at-risk type framework or derive real time measures of credit spreads in the mortgage market.

It also provides an anchor on what the GSEs should be charging in order to deliver a fair return to taxpayers and shareholders were the GSEs to be truly privatized again. I find that the market implied g-fee to be around 15 basis points on average, with a standard deviation of 6 basis points across different issuances. Mean implied g-fees are slightly higher for Group 2 CRT deals, those backed with collateral with higher loan-to-value (LTV) ratios. I also show that these results are robust to prepayment speeds and the assumption of losses following a beta distribution.

I find that implied g-fees are correlated to measures of credit spreads in other markets, which is intuitive given the systematic nature of mortgage credit risk. Some industry commentary has remarked that CRT bonds may reveal very little about the housing market because their floater spreads seem to co-vary strongly with credit spreads. But this may precisely mean that the market views housing risk similarly to corporate credit risk, and the portion of this risk premium co-moving with credit spreads could outweigh other variables like housing appreciation and delinquencies. This is in contrast to models such as Schwartz and Torous (1992), who argue a fair g-fee would load mostly on house price volatility and a systemic market risk premium for credit risk is not considered. This assumption appears to be strongly violated in the data.

With the loss distribution in hand, I analyze potential drivers of the "gap" between the average charged level of g-fees by the GSEs and the implied measure from CRT prices. After correcting for potential concerns about GSE retained risk, I calculate this gap to bee 16 basis points on average. This gap varies widely, in relationship to credit spreads as mentioned above, from more than 30 basis points to less than 5 basis points annually. Since we cannot observe directly the g-fees charged on the CRT collateral, I consider this an upper bound on the true gap.

I show suggestive evidence that the GSE retained portion of credit risk can explain some but likely not all of this gap. I also argue that it is unlikely this difference is due to moral hazard costs or CRT bond structure, since loan-level data on CRT loans is freely available, so any adverse selection would likely take place when placing loans into CRT deals, not afterwards (Echeverry (2020)). This leaves the explanation that this gap provides a measure of the economic cross-subsidization provided by the high quality loans underlying the CRT bond programs. This cross-subsidization occurs because the GSEs charge g-fees that are higher on high quality loans than the private market would charge in order to fund affordable housing programs, an important part of the GSE mandate (Cooperstein and Stegman (2019), Goodman et al. (2022)).

Estimates of this subsidization often focus on the default risk of high quality mortgages under the physical measure, and assume a constant return on equity for the GSEs. The strong co-movement of implied g-fees with credit spreads suggest that this may not tell the whole story. Time variation in the g-fee gap suggests that the economic value of this subsidy is changing over time. Put differently, when risk premiums are high, the GSEs are taking on more risk then econometric models would suggest, and they should be compensated for doing so. This also has the implication that they should either subsidize less low-quality loans during these times, or effectively take on more taxpayer risk (or shareholders, in the event the GSEs privatize). I do not argue this necessarily undesirable: home ownership can be a worthwhile policy objective. But without a market price for conforming credit risk, however, this gap can not be quantified. To my knowledge, this paper is the first to do so.

The remainder of the paper proceeds as follows: I briefly review the related literature below in section 1.1.1. In section 1.2, I describe the economics of the CRT program. Next, in section 1.3 I describe the data I will use in this paper before delving into a brief historical perspective on conforming mortgage defaults and losses in section 1.3.1. This historical perspective will provide the relevant benchmarks I can use to assess the results of the model and provide context. Section 1.4 explains the derivation and estimation of the credit model. In section 1.5 I discuss the estimation results, the implied distribution of mortgage losses, and the implications of the results for the GSEs and a calculation of the gap between charged and implied g-fees. Lastly, in section 1.6 I conclude.

### 1.1.1 Related Literature

This paper falls at the intersection of two main lines of literature; firstly, as the valuation method contained here employs a top-down approach, this paper contributes to the literature suggesting that top-down modeling provides accurate and parsimonious pricing models for pools of loans, and therefore MBS, CMOs, and as described here, CRT bonds. And more generally, the study of the GSEs, their mortgage guaranty, and their role in the housing market. While I focus on the provision and pricing of this guaranty, the CRT market offers a host of important economic questions that have yet to be addressed.

I combine a reduced-form credit model in which defaults are given exogenously specified with a top-down model in which mortgage cash flows are modeled at the aggregate pool level. Mortgages represented by CRT bonds are large groups of mortgages, between 50,000 and 180,000 similar individual loans per bond issuance, allowing the model to abstract from loan-level default behavior and retain the economic perspective on how defaults on a diversified pool of loans is priced. This treats mortgage default from a statistical perspective, much in the same way that a health insurer would estimate disease prevalence among its pool of insured persons rather than evaluate the personal risks for individuals. Giesecke, Goldberg, and Ding (2011) and others, for example, apply this technique to portfolio credit derivatives and Longstaff and Rajan (2008) apply this technique to CDO tranches to study how corporate credit defaults cluster. To my knowledge, I present the first use of this technique in order to model the paydown and default behavior of the mortgage pools backing CRT bonds. Fermanian (2013), Diener, Jarrow, and Protter (2012), and Sirignano and Giesecke (2019) provide further examples of this technique.

Early work on CRT bonds has been mostly qualitative and focused on discussing their potential role in alleviating capital requirements and credit risk at the GSEs. Wachter (2018) and Finkelstein, Strzodka, and Vickrey (2018) describe the role that CRT play in the economic platform of the GSEs. One recent quantitative study by Gete, Tsouderou, and Wachter (2020) study the response of CRT markets to Hurricane Harvey by utilizing secondary market spreads on the bonds. Gao and McConnell (2018) investigates the return on CRT tranches with respect to treasury bonds, albeit in a limited sample. Golding and Lucas (2020) simulated the returns to CRT bonds and finds that the mezzanine and junior tranches are subject to essentially no credit risk; I improve over this by noting that the risk-free values of mezzanine and senior tranches differ from their market values, meaning the market must believe that they are risky.

There is also a literature that looks at the level of the government subsidization of mortgage loans and the g-fee and its re-distributive role. This subsidization occurs both in the form of lower funding costs for the GSEs and an implicit "bail out" guarantee (Jeske, Krueger, and Mitman (2013)), or as cross-subsidization from high-quality to lower-quality borrowers (Gete and Zecchetto (2017)), or even across regions with differing local economic conditions (Hurst et al. (2016)). This literature is split on whether these subsidies help or hurt low income borrowers. My model provides a market based measure of the gap between charged g-fee levels, and so can speak to the amount of capital available for subsidization.

## 1.2 The GSEs, the CRT Market and the G-fee

Before discussing historical mortgage defaults and the valuation model, it is worth going through the history and current role of the GSEs and the economic structure of the CRT market. Prior to the CRT program, mortgage originators sold mortgages to the GSEs in return for MBS or cash, with the credit risk being retained on the GSEs balance sheets. US mortgage backed securities became a staple fixed income investment both in America and abroad, and contribute to lower mortgage rates.

This process of securitization has many benefits: for example, for homeowners, in the form of lower mortgage rates, and for investors, in the form of higher liquidity and more manageable risk profiles of MBS. The GSEs, in turn, must manage the credit risk of the mortgages that they purchase. This famously became a problem in the subprime mortgage crisis as the GSEs began to suffer large losses on loans that they had guaranteed. Due to the systemic nature of the mortgage market, the Federal Housing Finance Agency (FHFA), the GSEs regulator, placed the GSEs into government conservatorship on September 7th, 2008. They have remained there to this day, leading to significant unanswered questions about when, if ever, they would return to private ownership and how the systemic credit risk they help manage will be handled in the future.

One effort to mitigate the concentration of such a massive amount of risk in just two monoline entities was the creation of credit risk transfer (CRT) programs. Freddie Mac pioneered a particular form of the CRT program in 2013 when they issued the first bonds directly linked to the credit performance of loans they purchased and securitized. Fannie Mae quickly followed suit, and by the end of 2020 nearly half of all mortgages passing through the GSEs had their credit risk sold off to investors in the form of CRT bonds. Table 1.1 introduces a timeline of some of the main events in the history of the development of the CRT program.

| Date | Event |
|---|---|
| September 6, 2008 | The FHFA places both Fannie Mae and Freddie Mac into federal conservatorship in light of deteriorating housing market conditions. |
| September 2012 | The FHFA encourages the GSEs to develop risk sharing programs. |
| July - October 2013 | FNMA and FHLMC issue their first credit linked notes, known as fixed-severity deals, which feature predetermined write downs in the event of credit events. |
| January 2015 | CRT bond programs continue to develop, with actual loss deals taking over, in which CRT tranches are written down based on the actual losses realized by the GSEs on the loans. |
| July 2017 | Hurricane Harvey causes stress in CRT market. |
| Late 2018 | The first CRT REMIC designation occurs, making CRT bonds more tax effcient and attractive to investors. |
| March 2020 | Disruptions in the CRT market due to the COVID-19 pandemic: |
| July 2020 | Freddie Mac announces continuation of CRT bond programs after just a five month break during the crisis, quelling fears about CRTs long term viability. |
| October 2021 | Fannie Mae issues their first CRT bonds since the COVID-19 pandemic. |

**Table 1.1:** Timeline of Key CRT Market Events. The development of the CRT market has featured dynamic shifts in the level of credit protection provided, the tax treatment of the bonds, and more.

Credit risk transfer bonds are catastrophe bonds on the underlying mortgage collateral they represent. In this sense, they are a credit-derivative on conforming mortgages. When the GSEs purchase loans, a particular cohort may be selected to be linked to CRT bonds. The principal way in which securities are grouped is by placing loans with a $60 - 80$ loan-to-value ratio in a so-called Group 1 deal, and those with an $> 80 - 97$ LTV in a group 2 deal; both Fannie Mae and Freddie Mac follow this practice. In this paper, I continue this distinction by presenting and discussing the data and results at the GSE/LTV Group level.

CRT investors provide capital upfront (unlike a CDS contract); when losses are subsequently realized, the principal balance of these loans will be written down. These bonds protect a portion of credit risk faced by the GSEs and the mortgage loans they back. Only a limited amount of the credit losses can be attributed to write-downs of CRT bonds, which means that the GSEs will still be on the hook for losses realized above the amount of protection that the bonds offer. These features present a unique modeling challenge for extracting the informational content of these bonds and drawing conclusions about the implications for GSE credit risk management and the future of the US mortgage market. In order to address these challenges, I view each CRT tranche as partially revealing a portion of the loss density function. I specifically allow these probabilities to incorporate the possibility that losses exceed the detachment points of the particular tranche.

### 1.2.1 How does the CRT fit into the GSE market model?

Figure 1.3 shows the structure of risk transfer in relationship to the GSE and the underlying mortgages. Sitting at the center, the GSE purchases mortgages conforming to their guidelines and subsequently receives the g-fee for guaranteeing the credit risk on the mortgages. Why must the GSE reimburse the mortgage pool it owns for defaults? Precisely because the vast majority of GSE mortgages are securitized into credit guaranteed mortgage-backed securities. MBS investors, on the left of the chart, receive the cash flows from the underlying mortgages less the g-fee. If default occurs, the GSE reimburses itself from the CRT capital trust.

The right hand side introduces the CRT market. It is important to note that the CRT and MBS sides of the GSE business are separate. The MBS market does not need the CRT business to exist, and vice versa. Most importantly, cash flows from the mortgages are reserved for MBS investors and never flow to CRT investors.

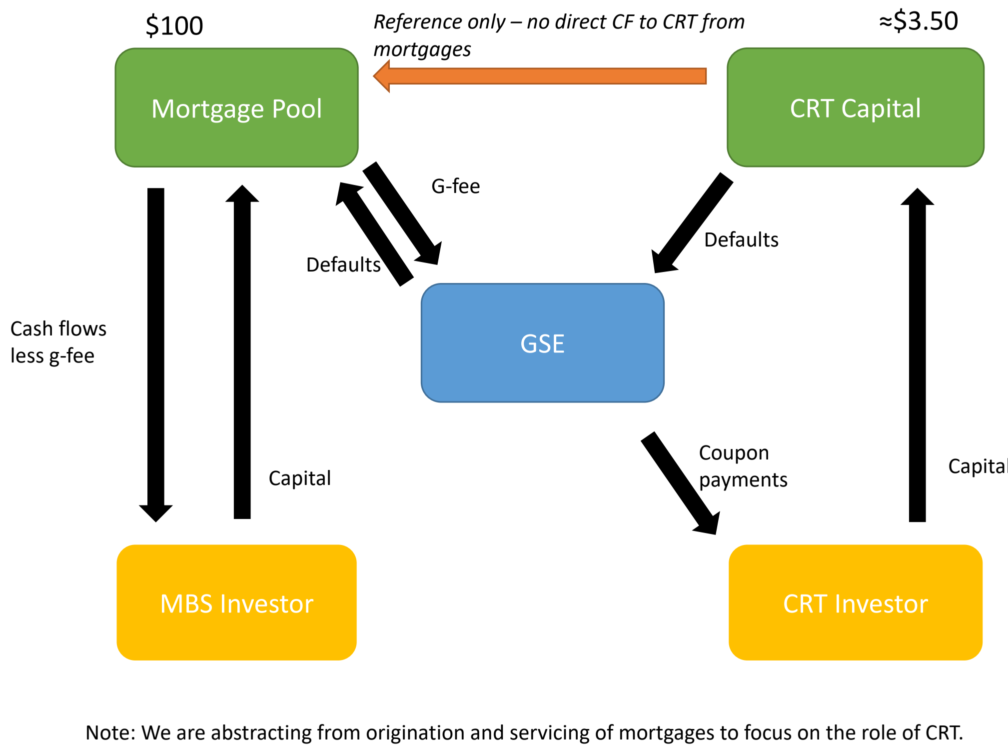



**Figure 1.3:** The typical CRT structure. Note that the CRT cash flows are obligations of the GSE and not the underlying mortgage pool. CRT bonds are distinct from MBS/CMOs in this sense.

It is in this sense that CRT are synthetic credit derivatives, whose cash flows are contractually linked to the performance of the underlying mortgage pool.

One can see immediately how the cash flows on CRT bonds mimic the g-fee, which passes through the GSE to the CRT investor in the form of coupon payments. Just as the GSE reimburses the MBS investors for losses, the CRT capital pool reimburses the GSE. But why is structure the way that it is? One can imagine many alternate structures, such as something akin to a credit default swap or just the sale of credit -risky MBS. In 1.2.2 below, I outline the economic arguments that lead to the development of the current CRT bond structure.

### 1.2.2 Economic Arguments

One argument for the development of CRTs in their current form is to avoid disruption of the highly liquid agency MBS market. The majority of agency MBS trading occurs in the to-be-announced (TBA) market; many have argued that the structure of this market benefits homeowners through lower mortgage costs.[7] An important founding principal of the CRT bond programs was that they would not interfere with this side of the business.

Secondly, it was important that CRT pools be fully funded so that the GSEs would not risk suffering losses on MBS backed loans while at the same time finding its CRT owners in distress and unable to pay, for example if CRT bonds mimicked a CDS contract. This would be a reasonable scenario in a time of crisis and would most likely amplify systemic risk, not quell it. Layton (2020) provides a deeper explanation of these and other principals that were factored into early CRT discussion and design.

### 1.2.3 The Structure of CRT Deals

Credit risk transfer deals are structured in a way that parallels the collateralized mortgage obligation (CMO) market, but differ in a few key dimensions. In this section, I outline exactly how a CRT deal is structured. Similar to CMOs, cash flows to investors are allocated to various tranches depending on their seniority. Typically, CRT deals involve between one and three mezzanine tranches (denoted by M: M1 would be most senior mezzanine trance, for example), and one or two junior tranches (called the B tranches). As the market has matured over the last 7 years, more tranches have been offered in general in order to appease the growing needs and desires of investors but the general structure has remained the same.

The most basic feature of a typical CRT bond is the allocation of principal pay-

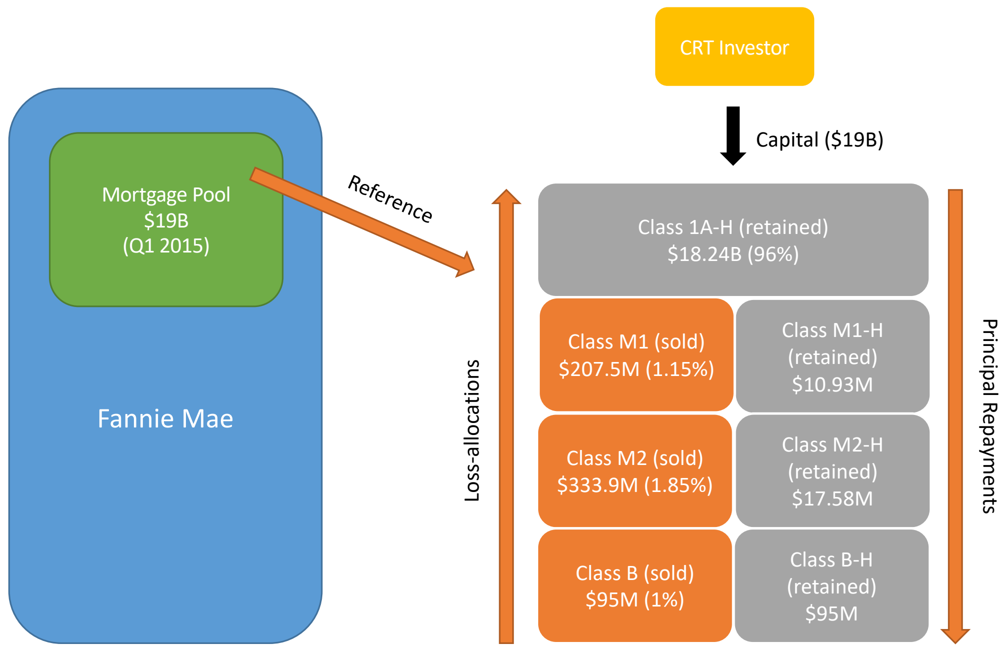
7. See Vickery and Wright (2013) for a discussion of this market.

ments and mortgage losses. Principal payments reduce the bond balances sequentially: the most senior bonds are paid back first, and the most junior bonds are paid off last. Losses are allocated to CRT tranche balances sequentially as well, but in the opposite order: starting with the most junior tranches first. Early bond issuances were called "fixed severity deals," because the language in these bond contracts specified certain bond write-down amounts in the event of distress in the underlying mortgage pools. Since 2015, both Fannie Mae and Freddie Mac have been issuing CRT deals under the so-called "actual-loss" framework. My model will abstract from these features, and a few other peculiarities of the bonds, because they would add vastly more modeling complexity with little added economic intuition. Since the fixed-loss write-downs were based on historical loss rates and default behavior, I simply assume that losses would be comparable to if they were actual-loss bonds on average. Recall that my model will simply extract the probability that tranches experience a principal write down.

#### 1.2.3.1 CRT Issuance Example

Figure 1.4 shows the issuance structure for the Fannie Mae CAS 2016 C-01 deal. This deal, issued in March 2016, features $19 billion of loans originated and acquired by Fannie Mae in early 2015. As mentioned above, it is important to note that although there is a CRT balance associated with all $19 billion of the loans, the majority of the balance is hypothetical. Only the notes sold to investors feature actual cash flows, where is the rest are hypothetical bonds used for accounting purposes. This will become more clear in the example.

This particular issuance features three notes sold to investors. Two mezzanine tranches, M1 and M2, and a junior tranche, B, which is exposed to the first-losses potentially realized in the reference pool. The initial credit enhancement in this deal is 4%: the B tranche is exposed the first 1% of losses and the mezzanine tranches the




Note: Example adopted from Fannie Mae CAS 2016-C01 Group 1 issuance term-sheet.

**Figure 1.4:** The typical CRT note structure. Note that the GSE retains a vertical slice of each tranche sold to investors - here, it is 5% of each mezzanine tranche and 50% of the first-loss piece.

next 1.85% and 1.15% respectively. Note that Fannie Mae retains a 5% vertical slice of each mezzanine tranche sold to investors, in order to maintain incentives (sometimes called the skin-in-the-game requirement). Fannie Mae also retains half of the first loss piece and all of the senior $18.24 billion portion. Appendix section 1.A.4 shows a simple example of cash flows would be allocated to a generic CRT style bond. The impact that this risk retention could have on pricing and adverse selection problems if it were to be removed is beyond the scope of this paper, but would be interesting avenue for future research.

## 1.3 CRT Data and Summary Statistics

The main data source for the project is a hand collected data set detailing 121 CRT bond issuances starting with the first CRT deal, issued by Freddie Mac in July 2013, and ending with the latest CRT issuance in January 2022. Only “traditional” CRT bond issuances from the GSEs are included, as the modeling of seasoned and more bespoke CRT transactions are beyond the scope of this project.[8] Traditional bond CRT issuances are those backed by a reference pool of newly originated and acquired mortgages upon which claims to principal are sold in structured securities. As described in the previous section, the hallmark characteristics of these transactions are the distinction between the retained senior portion and the subordinated portion, with principal returned pro-rata between the two and sequentially among subordinated tranches and losses allocated in the opposite direction.

To build the data set, CRT deal information is hand collected from the GSE websites through prospectuses for the CRT bond deals. These prospectuses contain information on bond level information such as total principal amount, the tranches offered and their CUSIP numbers, attachment/detachment points, and expected credit ratings. Table 1.2 below shows the number of bonds issued for each GSE/Group level deal. To be consistent with how the bonds are described by both the GSEs and the popular press, we bucket each issuance based on the issuing GSE and the LTV bucket. Recall that the group number refers to the LTV of the underlying loans, with Group 1 issuances referring to underlying LTVs of 61-80, and Group 2 referring to the ”high LTV” deals of 81-100. There are a total of 400 individual bonds that comprise the data set:

The average deal size for offered bonds ranges from $700 million to around $1

8. The GSEs have experimented with some unique bond issuances over the life of the CRT program. These include so-called ”seasoned B” transactions, which feature previously retained B tranches, among others. See Freddie Mac FTR transactions for example.

| Group | Bonds Issued | Attach | Detach |
|---|---|---|---|
| FHLMC-G1 | 138 | 0.17 | 3.91 |
| FHLMC-G2 | 110 | 0.19 | 4.71 |
| FNMA-G1 | 90 | 0.33 | 3.87 |
| FNMA-G2 | 62 | 0.34 | 4.09 |

**Table 1.2:** Summary Statistics by GSE Issuance Groups. The average attachment and detachement points show the typical credit coverage that the CRT deals provide, between 25 and 400 basis points on average.

billion, implying that the total notional of the underlying mortgage loans is around \$20-35 billion on average if we assume around 4% credit coverage on average. In order to supplement this data set with pool-level characteristics, I combine the bond level dataset with data on the underlying mortgage pools. The end result is a dataset that combines two levels of information - “deal level” statistics such as weighted average pool coupons and credit scores, and “bond level” information such as individual tranche attachment and detachment points. The dataset is fed into the pricing model at the deal level in order to estimate the deal level default probabilities.

Lastly, the TRACE agency data set is leveraged in order to match market prices during the first week of trading. The market price is taken as the weekly average trade price following issuance. This is done in order to gauge the market view of credit risk in the pool before any substantial news regarding defaults in the pool could be learned. Table 1.3 below goes into further detail, showing average prices for each generic tranche type. We see that prices can move substantially away from their issuance price, emphasizing the the importance of using market prices rather than issuance prices in the fitting of the pricing model.

For graphical exposition, Figure 1.5 shows key time-series summary statistics for the underlying mortgage pools. We see that the weighted average maturity of underlying loans has increased but only slightly, reflecting that loans are packaged into the CRT pipeline on average around 7 months after origination. LTV ratios are flat, corresponding the group type of the particular deal that they represent. Weighted

| Group | Tranche | N | First Issued | Last | Spread | Min Price | Mean P | Max P |
|---|---|---|---|---|---|---|---|---|
| FHLMC-G1 | B1 | 23 | 2017-02-07 | 2022-01-21 | 378.3 | 98.93 | 100.51 | 101.60 |
| FHLMC-G1 | B2 | 28 | 2015-02-03 | 2022-01-21 | 835.9 | 91.31 | 100.88 | 108.84 |
| FHLMC-G1 | M1 | 38 | 2013-07-26 | 2022-01-21 | 107.8 | 99.89 | 100.10 | 100.87 |
| FHLMC-G1 | M2 | 37 | 2013-07-26 | 2022-01-21 | 246.6 | 99.30 | 100.38 | 104.01 |
| FHLMC-G1 | M3 | 12 | 2014-02-12 | 2016-09-30 | 430.8 | 99.81 | 100.81 | 102.52 |
| FHLMC-G2 | B1 | 18 | 2017-02-22 | 2021-12-10 | 399.7 | 95.17 | 100.10 | 101.73 |
| FHLMC-G2 | B2 | 23 | 2015-03-31 | 2021-12-10 | 889.3 | 98.92 | 101.30 | 106.21 |
| FHLMC-G2 | M1 | 29 | 2014-08-11 | 2021-12-10 | 103.3 | 97.86 | 100.01 | 100.52 |
| FHLMC-G2 | M2 | 29 | 2014-08-11 | 2021-12-10 | 238.4 | 97.12 | 100.12 | 102.50 |
| FHLMC-G2 | M3 | 11 | 2014-08-11 | 2016-10-25 | 440.0 | 98.93 | 100.53 | 102.09 |
| FNMA-G1 | B1 | 23 | 2016-02-18 | 2022-01-20 | 544.6 | 99.83 | 101.47 | 106.50 |
| FNMA-G1 | B2 | 3 | 2021-10-27 | 2022-01-20 | 583.3 | 100.31 | 100.38 | 100.50 |
| FNMA-G1 | M1 | 32 | 2013-10-24 | 2022-01-20 | 113.5 | 99.14 | 100.09 | 101.07 |
| FNMA-G1 | M2 | 32 | 2013-10-24 | 2022-01-20 | 332.3 | 97.03 | 100.69 | 105.88 |
| FNMA-G2 | B1 | 15 | 2016-04-21 | 2021-12-01 | 564.7 | 99.56 | 100.98 | 106.24 |
| FNMA-G2 | B2 | 1 | 2021-12-01 | 2021-12-01 | 620.0 | 101.57 | 101.57 | 101.57 |
| FNMA-G2 | M1 | 23 | 2014-05-28 | 2021-12-01 | 115.0 | 99.90 | 100.08 | 100.34 |
| FNMA-G2 | M2 | 23 | 2014-05-28 | 2021-12-01 | 350.2 | 97.19 | 100.63 | 102.97 |

**Table 1.3:** Summary Statistics by Generic Tranche. M bonds with lower numbers are more senior mezzanine tranches, and in contrast, B bonds with higher numbers are more junior and take the first loss. Hence, the table shows that they pay the highest spread.

average coupons follow mortgage rates for newly issued loans. Lastly, credit scores have declined slightly over the CRT program. This is attributable for increased interest for risk as the CRT program gained in maturity and popularity; the GSEs began to include more loans to meet demand.

Continuing with figure 1.6, we observe the attachment and detachment points over the last 9 years of bond issuances. Early CRT deals did not feature first-loss components, which both GSEs began to feature in 2016 going forward. The tranches offered can be seen as a function of investor appetite as well as the economic appeal offered to the GSEs of selling off that portion of risk. I attempt to model the paydown of the underlying loans independently of the tranches that are offered, but earlier issuances where less variety of bonds are offered may potentially be modeled less accurately. Future research should aim to further make the relationship between the structure of the CRT deal and the information content of bond prices independent.

Lastly, figure 1.7 shows the floater spreads at issuance for each generic tranche for

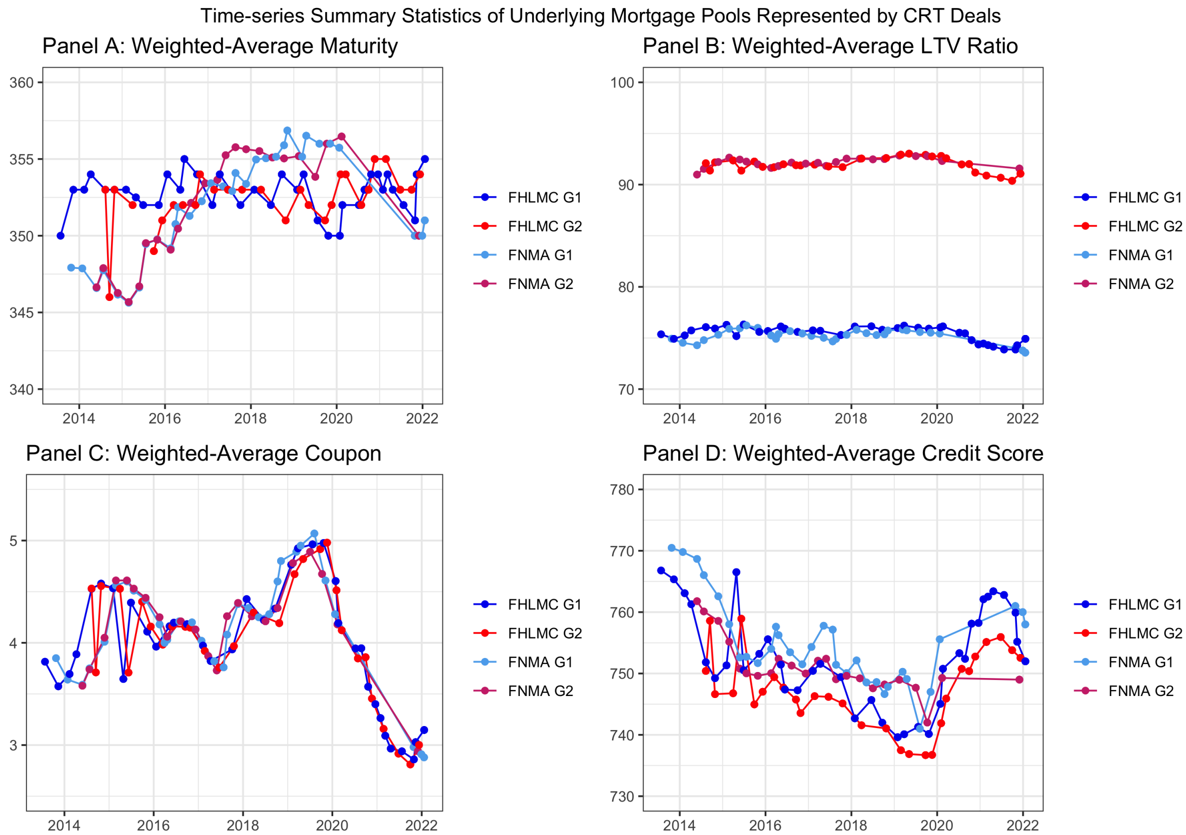



**Figure 1.5:** This figure plots summary statistics in the time-series for the characteristics of the mortgages underlying each CRT bond deal. Some interesting trends emerge; in particular, the weighted maturity of CRT deals has increased over time indicating that the pipeline for origination to credit protection has decreased. Also, the typical credit score has decreased over time, suggesting that riskier mortgages are being included in the CRT deals. See text for more details.

both the CAS and STACRS notes. Of course, the riskier notes pay higher floaters spreads. These spreads can go through periods of relatively stability, but are also clearly affected my market wide credit risk as can be seen when CAS spreads spiked in early 2016 and when Freddie Mac issued bonds during the COVID-19 pandemic in mid 2020.

The aim of this section was to provide a comprehensive view of how the CRT market is structured and how it has evolved over time. One can see quite clearly that there is substantial heterogeneity in CRT bond issuance over time in terms

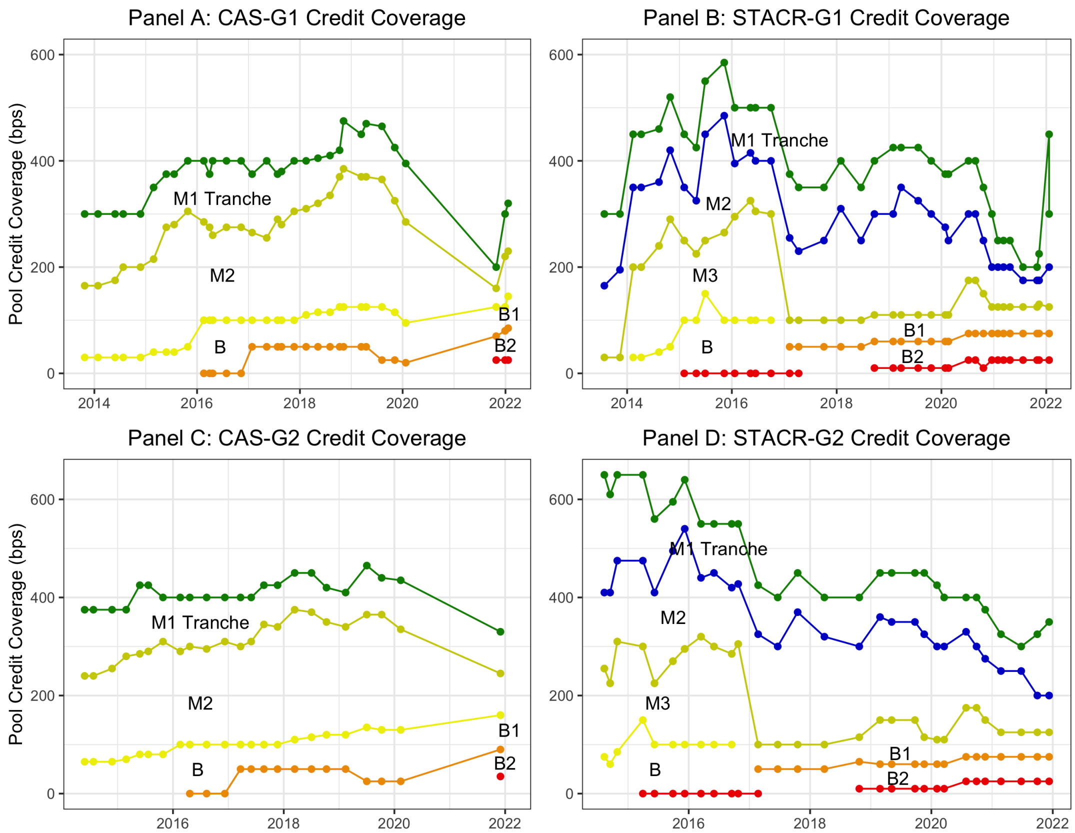



**Figure 1.6:** The time-series of Fannie Mae CAS note attachment and detachment points for Group 1 and Group 2 issuances. Lines represent breakpoints between different bond types; for example, in 2015, M2 note holders started taking losses when collateral losses reached 50 basis points. Losses then accrued to the M1 note holders as the credit protection by the M2 notes was exhausted around 200 basis points. This plot demonstrates the changing nature of credit protection and its evolution over time.

of bonds offered, but it is also important to note that the underlying loan quality has been relatively constant throughout time, making the modeling choices made in this paper reasonable ones. While the mortgage market always contains uncertainty regarding underwriting quality, most commentary concludes that the period during which CRT bonds have been offered since 2013 has been relatively consistent and of high quality. There is also not much concern that loans differ substantially between the bonds backing Freddie Mac and Fannie Mae CRT programs. In fact, the loans originated between the two are substantially similar enough that the MBS market

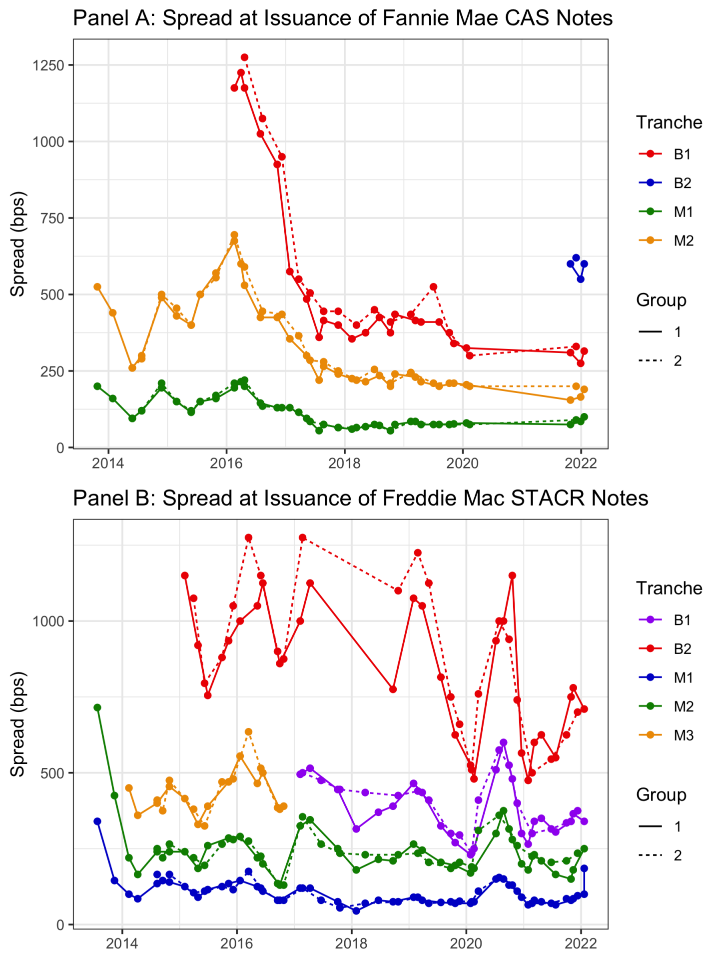



**Figure 1.7:** CRT Bond spreads to LIBOR at issuance for both Fannie Mae CAS and Freddie MAC STACR notes.

was consolidated in 2019 to offer a single security for which loan pools from either GSE are acceptable for TBA delivery to the purchaser of an MBS. In the next section 1.3.1, I briefly describe historical defaults in conforming mortgage pools.

### 1.3.1 Mortgage Losses: A Historical Perspective

As a segue before describing the valuation model and estimation procedure, I will briefly examine the historical performance of mortgage cohorts for conforming loans purchased by the GSEs. Historical defaults, especially during the crisis, provide some of the justification for why CRT subordination levels fall around 4%. They also help fix ideas about the types of risks that mortgage credit investors may be pricing (as in O'Neill (2022)).

Figure 1.8 utilizes the Fannie Mae Single Family Data set, a comprehensive dataset released by Fannie Mae that covers over 50 million single family loans originated since mid-1999.[9] It shows historical losses for cohorts of loans originated in a particular year and sold to Fannie Mae. The left-hand panel shows all years in the single-family data set, and the right-hand panel omits the years 2004-2008. Loans originated just before and during the crisis suffered total losses peaking at just over 350 basis points of total cohort notional. Loans originated during more tranquil times, around the start of the FNMA data set, ended up with losses around 20-25 basis points or so. Amortized using the 5 year average life of a mortgage loan, the "expected loss" portion of the g-fee is about 5 basis points for year. This back-of-the-envelope calculation shows where most industry/practitioner work is getting this number. One could also imagine at current g-fee levels, the catostrophic losses of upwards of 300 basis points may be priced as happening about $\approx 25/300 \approx 8\%$ of the time. Of course, investors in bonds will require a risk premium in the event that losses are higher, say on the order of the crisis period originations. My model can speak to the probability (risk-neutral) that investors prescribe to this state of the world.

In a follow up paper (O'Neill (2022)), I go into further detail on loss severity and time between default and loss in this data set in order to motivate a richer CRT bond model driven by two risk factors, one that represents baseline levels of mortgage default and one that seems related to disaster risk. Additionally, the inclusion of cohorts that suffered heavy losses during the subprime mortgage crisis provides evidence of the nature of the disaster states that investors in CRT bonds may be particularly worried about. They may require a large risk premium to provide insurance against

9. This dataset is extensive and is available publicly through Fannie Mae Data Dynamics.

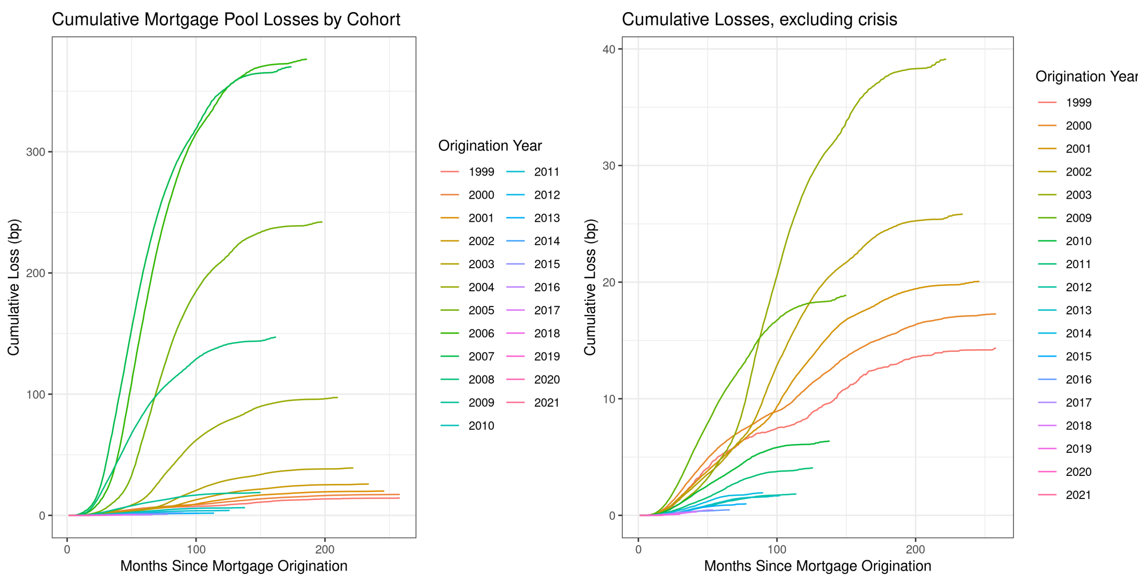



**Figure 1.8:** Cumulative losses for historical FNMA mortgage cohorts. These cohorts are filtered to include only loans that have LTVs between 60 and 97 and credit scores between 640 and 780, to make them most comparable to the CRT collateral.

these states.

## 1.4 Valuation Framework

For consistency, this section describes the valuation procedure in the context of the stylized CRT issuance that was described above. Recall that this example CRT issuance has three tranches: a junior, first-loss piece, a mezzanine tranche (often called the M2 tranche as described above), and the senior M1 piece. For a generic tranche $j$, I define the attachment/detachment points as $K_j$ and $K_{j+1}$ respectively. Valuation proceeds in 3 main steps, which emphasize the intuition that CRT tranches are derivatives of the same underlying loans and therefore the paydown of those loans is the primitive modeling object. The senior and subordinate portions are in turn claims on the paydown of the loans, and lastly the CRT bond cash flows are functions of the balances of the subordinate portion.

The main intuition behind the valuation model is that the difference between the hypothetical tranche price with no defaults (the risk-free price) and the market price is the present value of losses that the market expects the bond will incur. I will go into further detail in this section at how I arrive at these risk-free prices, as well as the particular modeling quirks associated with the paydown of the mortgage pool.

### 1.4.1 Cash-flow Model

In order to calculate the value of the CRT tranches in both the risk-free and various default scenarios, I start by describing the paydown of the underlying loans. I assume that the underlying loan pool consists of loans that are homogeneous in their interest rate, $r$ (the weighted average coupon of the loan pool), and maturity $T$, which I hold to be 360 as in a typical 30 year loan (in reality, a small adjustment must be made to account for the fact that the loans are a few months into their term at the time of bond issuance). The total starting principal balance is normalized to equal 1. This homogeneity is well supported assumption in the data, as similar to MBS, loans packaged into CRT deals are originated at around the same time and the majority fall into a 50 basis point rate bucket. I believe the simplicity of this modeling approach outweighs the potential downsides from slight rounding errors in principal payments.

In order to properly keep track of the principal payments, both scheduled and unscheduled, occurring at a given time, we must distinguish between the notional balance remaining in the pool (upon which mortgage payments and prepayments are based), and the actual pool principal balance, which is reduced overtime as loans are paid back or defaulted upon, which I define here as $B_t^*$.

A key modeling object in these types of models is the notional loan amount that has yet to be prepaid at a given time, which I define as $P_t$. $P_t$ can be expressed as the cumulative product of single month mortalities ($SMM_t$), the percent of loans that are prepaid in a given period.

$$P_t = \prod_{n=1}^{t} 1 - SMM_n \tag{1.2}$$

If this was an MBS model, we could stop here and define scheduled and unscheduled principal payments in terms of this factor. Since CRT tranches contain credit risk, by definition, I introduce a new layer to this calculation and similarly define the object $D_t$, which represents the proportion of principal that has yet to default at time $t$. The default corollary to the SMM is termed the monthly default rate, MDR. Combining these two terms, we can express the total notional principal at time $t$ as the product of these two objects, $Q_t = P_t \times D_t$, called the survival factor. Thus, the total principal remaining in the pool at a given time $t$, $B_t^*$, is given by:

$$B_t^* = P_t \times D_t \times B_t = Q_t \times B_t \tag{1.3}$$

Here, $B_t$ would be the principal balance at time $t$ in the absence of prepayments or defaults. This is readily seen in the case where $SMM$ or $MDR$ are equal to 0 for all $t$. For brevity, I do not include the all of the mortgage paydown formulas here as they are standard once incorporating a new, default sensitive survival factor $Q_t$ as I have defined above.[10] For example, scheduled principal in period $t$ would be the scheduled principal in the no-default, no-prepayment scenario multiplied by the survival factor at the beginning of the period, $Prin_t^* = Prin_t \times Q_{t-1}$.

Lastly, in the reduced-form credit model, the losses realized on the mortgage pool, $L_t$, are driven by an exogenous process. I will specify in the following section the exact functional form that losses will be allowed to take in this particular model, but in this general framework, loss amounts (the corresponding MDR) would be given by $l_t = MDR_t \times (B_{t-1}^* - Prin_t^*)$. Cumulative losses would simply be $L_t = \sum_{n=1}^{t} l_n$.

The second step in the modeling process is to specify the paydown of the senior and subordinate portions of the corresponding CRT deal. Recall that in a typical deal, the first 4% or so of losses is sold to investors and corresponds to the subordinate portion of the principal balance. The remaining is the senior portion, which is retained by the GSEs themselves. Scheduled and unscheduled principal is paid down pro-rata between these two groups.

Note, that this subordination level can change over time as losses are written down on the subordinate balance. The expression for the senior balance over time is intuitively made up of its pro-rata claim on prepayments and scheduled principal payments less its claims on losses that exceed the subordinate balance. Here, $Ppmt^*$ and $Prin^*$ are period $t$'s scheduled principal and prepayment amounts.

Assuming normalized starting balances $Sen_t = Sub_t = 0$ for $t = 0$:

10. Hayre (2001) contains a summary of these different formulas in the context of a pass-through mortgage backed security.

$$Sen_{t+1} = Sen_t - \underbrace{\frac{Sen_t}{B_t^*}}_{\substack{\text{Pro-rata} \\ \text{portion}}} \times \underbrace{[Ppmt_t^* + Prin_t^*]}_{\substack{\text{Scheduled Principal} \\ \text{and Prepayments}}} - \underbrace{\mathbb{1}_{Sub_t=0}[l_t]}_{\text{Loss Claim}} \tag{1.4}$$

The last term represents the losses that have accrued to the senior claim if the subordinate claim has already been entirely written down. And analogously for the subordinate balance, which has the slight difference that once it is written down to zero, it cannot turn negative as loss claims have now shifted to the senior portion:

$$Sub_{t+1} = \Big[Sub_t - \underbrace{\frac{Sub_t}{B_t^*}}_{\substack{\text{Pro-rata} \\ \text{portion}}} \times \underbrace{[Ppmt_t^* + Prin_t^*]}_{\substack{\text{Scheduled Principal} \\ \text{and Prepayments}}} - \underbrace{\mathbb{1}_{Sub_t>0}[l_t]}_{\text{Loss Claim}}\Big]^+ \tag{1.5}$$

The last component of the valuation expression is to model the CRT tranches as functions of the underlying subordinate balance. CRT tranche balances can be thought of as call spreads on the subordinate balance and cumulative mortgage losses. The paydown and default dynamics of the underlying mortgage pool lead to the amortization of both the senior and subordinate claims to the mortgage cash flows; in addition, the default losses incurred on the mortgages are accrued first to the subordinate balance and then to the senior balance if losses exceed the amount of credit protection offered by the subordinate balance.

The balance of a particular CRT tranche can be thought of as a collection of call spreads. If the principal balances of CRT bonds were not dynamically paid down as loans were paid off, the expression would be simple. However, since principal from the loans flow in over time, the effective amount of credit risk each tranche is taking is reduced as well. The cumulative loss claim at time $t$ for tranche $j$ is equal to:

$$LC_t = [L_t - K_j]^+ - [L_t - K_{j+1}]^+ \tag{1.6}$$

We recognize this as being long a call spread on losses with strikes of $K_j$ and $K_{j+1}$. Analogously, the cumulative principal claim at time $t$ for tranche $j$ is the following, where $\eta$ represents the initial credit enhancement provided by the subordinate balance:

$$PC_t = [Sub_t + K_{j+1} - \eta]^+ - [Sub_t + K_j - \eta + LC_t]^+ \tag{1.7}$$

Putting all of the pieces together, we have an expression for the balance of tranche $j$ at time $t$, which we denote by $B_{K_j,t}$. We scale by the size of tranche $j$ relative to the overall notional value of the mortgage pool in order to scale the notional of tranche $j$ as equal to \$1:

$$B_{Kj} = \frac{[K_{j+1} - K_j - PC_t - LC_t]^+}{K_{j+1} - K_j} \tag{1.8}$$

Plugging in and reducing:

$$B_{K_{j,t}} = \left[1 - \frac{[Sub_t + K_{j+1} - \eta - LC_t]^+ - [Sub_t + K_j - \eta - LC_t]^+}{K_{j+1} - K_j} - \dots \right. \\ \left. \dots \frac{[L_t - K_j]^+ - [L_t - K_{j+1}]^+}{K_{j+1} - K_j}\right]^+ \tag{1.9}$$

The value of the tranche is thus given by the sum of present value of the coupon payments paid at floater spread $s_{K_j}$ above LIBOR/SOFR rate $q_t$ and the present value of the returned principal payments. Furthermore, define $D(t)$ as the price of a zero-coupon bond that pays \$1 at time $t$. The value of tranche $j$ is equal to the following, where the expectation term simply reflects the fact that we have not specified a process for prepayments or defaults, and these could indeed be random, and the time 0 principal starting at its normalized value of 1:

$$P(K_j, T) = \sum_{t=1}^{T-1} D(t) \times E_t^Q[\underbrace{B_{K_{j,t-1}}(q_t + s_{K_j})}_{\text{Coupon Payment}} + \underbrace{PC_t - PC_{t-1}}_{\text{Principal Repayment}}] + \underbrace{D(T)B_{K_{j,T}}}_{\text{Principal at Maturity}} \tag{1.10}$$

With these expressions in hand, the last four requirements for the valuation procedure are specification of prepayment rates, LIBOR/SOFR rates, the risk-free discount curve for the bond cash flows, and the default process. For prepayment rates, consistent with the goal of keeping the model as simple as possible, I choose a constant prepayment rate that ramps up over the first 30 months of the loan life (this is known as the PSA prepayment assumption and is commonly used in mortgage-backed securities modeling). This assumption could easily be altered to include random prepayments.

Similarly, for LIBOR rates (and later in the sample, SOFR rates), I choose to assume that the LIBOR rate is held constant over the life of the loan at the starting rate (i.e. the rate at the time of bond issuance and which will be paid on the bonds first coupon payment). Lastly, the discount function, $D(t)$ represents the present value of receiving one dollar at time $t$. Since this is a price of a risk-free zero coupon bond maturing at time $t$, I use the cubic spline approach to bootstrap the zero-coupon bond prices for the necessary maturities as in Longstaff, Mithal, and Neis (2005) from the yield curve of constant maturity treasury prices. The next section describes the default process.

### 1.4.2 Solving for Default Probabilities

To this point, I have not yet described the form that losses will take in this model. For simplicity, I take the approach of assuming that tranches are defaulted upon entirely in the event that a default state is reached. This simplification vastly reduces the computational cost of matching tranche values to market prices because tranche values can be readily solved for in the given default state. The intuition of the resulting model is that the difference between the market prices of CRT tranches and their default-free value discounted at the risk-free rate is equal to the present value of the losses on that particular tranche. Of course, because of the sequential paydown structure of the CRT bonds, tranches are not necessarily worthless even if their balance is entirely written down at some future default time $\tau$. Since I am implicitly assuming that there are $N$ total default states where $N$ is the number of tranches in a CRT deal, I must value the tranche in each scenario.

The starting point for the analysis is to propose the $N$ default states and default times. I take the default states to result in the write-down of the $N$ tranches and the default times to occur in the period in which amortization of that bonds principal balance would occur. Because of the sequential nature of the principal repayment, the write-down of a particular tranche results in the necessary write-down of all tranches its junior. In the case of the stylized CRT issuance we have been following along with so far, this results in the following loss scenarios:

- $a$: In scenario $a$, the $M1$ tranche, the most senior, tranche is written down at default time $\tau_a$. This implicitly means any tranches junior to $M1$ are written down as well.[11]
- $b$: $M2$, the mezzanine tranche, tranche is wiped out at default time $\tau_b$
- $c$: means that at least $B$, the most junior tranche , is written down at default time $\tau_c$

Now we must describe how we can map from market prices to the default probabilities associated with them. Figure 1.9 shows a tree that explains the model and the probabilities associated with each state. Since we have $N$ market prices, we can

11. For the most senior tranche, I take the default time to equal $t = 3$ instead of $t = 0$, since default at the very beginning of the pool paydown is implausible. Due to the low likelihood of default on the most senior tranche, this assumption has little to no effect on estimated loss distributions or g-fees.

solve for that many probabilities, one each associated with the write-down of that particular tranche.

Since the most senior tranche is the first to be paid back, I treat the probability of it reaching the default state as $a$. If the bond survives this default state, it is paid back and we reach the beginning of the amortization stage for bond $B$, which defaults with probability $b$ or time $\tau_0$ probability $(1-a)b$, and so on.

Define tranche $i$'s market price as $P_i$. We can also calculate the market price under the possibility of no loss write-downs, represented by $P_i^*$. Lastly, we consider $V_i$, to be the difference between the risk-free price and the market price:

$$V_i = P_i^* - P_i$$

In this framework, $V_i$ represents the present value of expected losses on a particular tranche $i$. This object contains information related to the probabilities of the particular losses that could be realized on the tranche. To fix this idea, consider the most simple security, one that pays \$1 tomorrow, or nothing. If the market price is

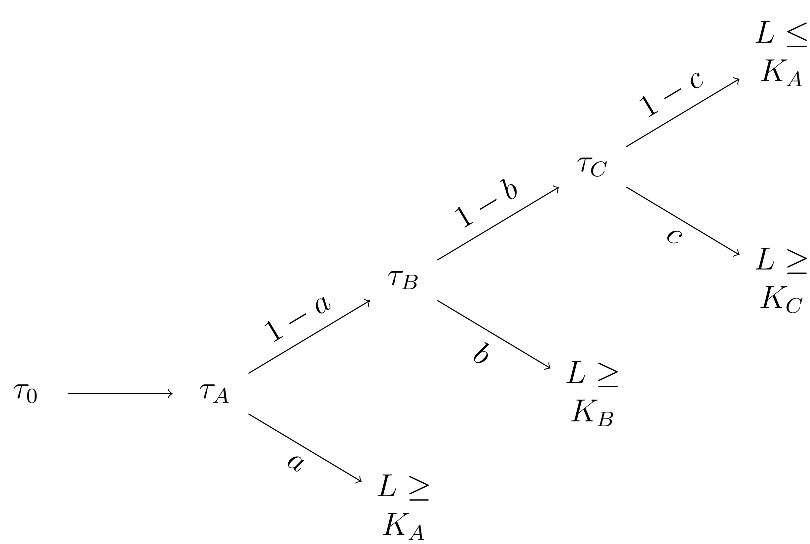



**Figure 1.9:** Tree model that shows the loss scenarios.

\$0.50, and the risk-free rate at 0%, the market is pricing a risk-neutral probability of receiving nothing at 50%. In our stylized model, we extend this same logic to the multi-dimensional case where there are several related securities and several possible events. Since there are 3 securities. I am able to calculate probabilities of default in 3 separate scenarios.

These probabilities lead to the following system of equations, where $V_{i,s}$ is the present value of losses in scenario $s$ for tranche $i$. The system is non-linear in the probabilities, but is easily solvable recursively, or by defining new variables in the form of $z = b(1-a)$ and solving in matrix form:

$$\begin{aligned} V_{M1} &= aV_{M1,a} \\ V_{M2} &= aV_{M2,a} + b(1-a)V_{M2,b} \\ V_B &= aV_{B,a} + b(1-a)V_{B,b} + c(1-b)(1-a)V_{B,b} \end{aligned} \tag{1.11}$$

I do not assume that CRT issuances can only suffer losses up to the amount of protection provided by the CRT bond issuance. Instead, I assume that the probabilities produced by the model are informative about losses that exceed the level of protection. For example, if there is a 60% chance that the most junior, first loss tranche is exhausted, that probability simply encodes the possibility that losses equal or exceed the detachment point of that tranche. The lack of market prices for tranches senior to that piece in that scenario preclude me from understanding losses above that point until I put further structure on the distribution in the next section.

Thus, the problem is reduced to using the valuation model described above, where the system can easily be extended to account for CRT issuances with more or less tranches. Of course, we can only price the number of scenarios for which we have tranche market price. Since probabilities sum to 1, we have that $\prod_{i=a}^{c} 1 - i$ equals the probability that losses below the subordinate attachment point have occurred. If this attachment point is zero, I am assuming that there are no losses on the pool.

### 1.4.3 Estimation

I estimate the model above to derive the associated probabilities of tranche defaults for the 121 bond issuances and 400 tranches in the data set. As alluded to in the previous sections, CRT bonds only offer snapshots into the incomplete probability density function of portfolio losses. In order to complete this picture, we must specify a distribution and estimate the parameters of said distribution. With the number of tranches in a given CRT deal ranging from 2-5, I opt for the Beta distribution, since its simple two parameter specification results in the moments being exactly identified in the less common 2 tranche case and over identified in the much more common 3, 4 and 5 tranche case. Furthermore, the beta distribution has the advantage of being bounded on the interval [0, 1], which makes it convenient for estimating portfolio losses.[12]

For a given CRT deal, I can estimate the risk neutral density associated with a loss exceeding each of the $N$ detachment points included in the deal. For example, the probability $a$ estimated in the model above is the probability that losses exceed the detachment point of the senior tranche, or $a = A = P(L \geq K_a)$. The probability $B = a + a(1 - b)$ is the probability that losses exceed the detachment point of the second-most senior tranche, or $P(L \geq K_b)$, and so on. The plots below in figure 1.10 plot this graphically for a generic beta distribution. Each probability estimated in the above model is used as a moment in order to fit the beta distribution associated with that particular deal issuance. The resulting beta distribution means that the market implied distribution $L \sim Beta(\alpha, \beta)$.

For each issuance, I fit $\alpha$ and $\beta$ of the beta distribution parameters that best match the moments implied by the market prices as described above. I scale the Beta

12. Using this particular distribution is a modeling choice, and although I find that it works well to describe the probabilities generated by the pricing model, future researchers may prefer a different distribution with at most $N$ parameters and preferably less.

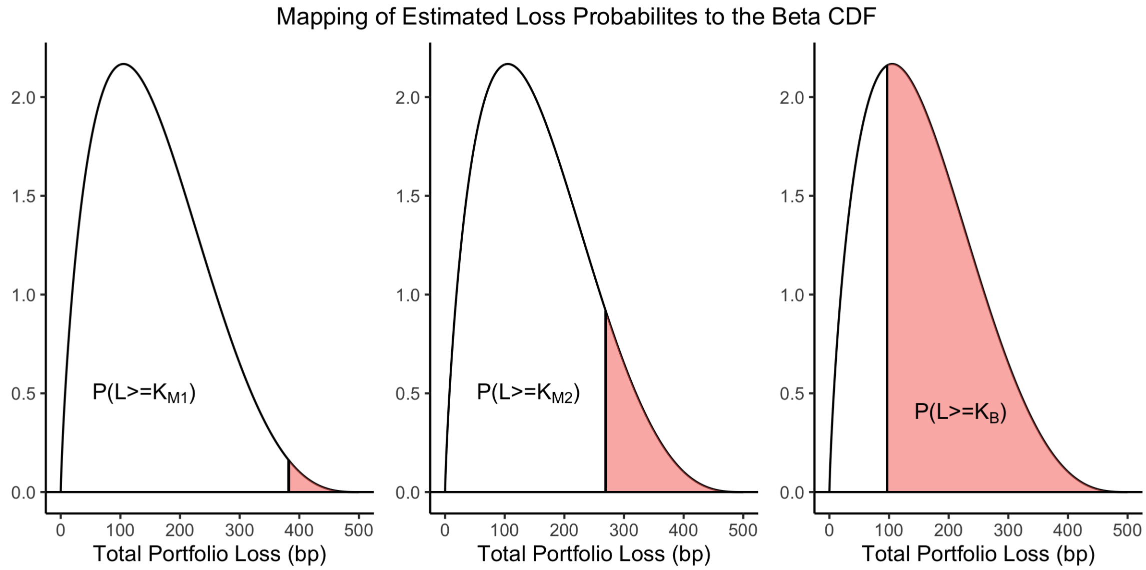



**Figure 1.10:** Example distribution from the model described in the text. The probabilities encapsulated in the present value of losses on each tranche map to probabilities in the risk-neutral loss probability density. See the text for more details on how these probabilities are derived.

distribution to have the support of most senior attachment point plus an additional 25 basis points. This effectively places a cap on the maximum loss the pool could sustain, but since so little of the probability density is in the tail, this choice has little effect on the expected value of the distribution, which is how I will map from this distribution to a market implied g-fee.

## 1.5 Results and Implications for the GSEs

In this section we first discuss the results of the valuation and estimation procedure, and then discuss implications for the GSEs and future research. To reiterate, one beta distribution is fit for the 121 bond issuances in the CRT data set using the probability moments (that is, the piece of each implied loss function probability density function) derived from market prices using the model in section 1.4.2.

The fitted distributions as described here are fitted using a conditional prepayment rate of 15%.[13] The beta distribution fits the cumulative probabilities very well in cases where we have more tranche prices than parameters (since we of course can fit the distribution exactly in the CAS issuances in which there are only 2 bonds). The RMSE expressed in percentage terms is around 25 basis points. For comparison purposes, the mean value of $a$, the probability the most senior tranche is written-down, is 1.3%. Because the other probabilities themselves correspond to tranches that can vary quite a bit in terms of attachment and detachment points, it does not make much sense to look at the time-series of probabilities themselves.

### 1.5.1 Fitted Loss Distributions

Figure 1.11 shows the risk-neutral beta distributions for a sample of bond issuances from different years and GSE issuers. It is immediately apparent that the shape of the distribution varies over time, with tail risk heightening in 2016 as credit spreads widened during surprising political elections in the US and UK and other world events.

What else can we learn from these distributions? They provide a unique measure of real-time risk that is much more transparent than those provided by more complicated models. Research from mortgage-backed securities markets, for example, shows that option-adjusted spread models can be heavily model dependent and may say more about the underlying model then about the security they are pricing.[14] My model has the advantage of deriving an arbitrage-free estimate of the implied loss-distribution that is not dependent on a particular econometric model of defaults, but rather one that is derived directly from tranche market prices. They could also be

13. Appendix 1.A.7 shows that the implied g-fees from the baseline model are not very sensitive to the prepayment rate, although the implied beta distribution will change depending on if present values in default scenarios have increased or decreased.

14. Diep, Eisfeldt, and Richardson (2021) show MBS OAS vary widely across dealers for the same security and Chernov, Dunn, and Longstaff (2017) Figure 1 shows substantial movement of OAS upon model updates.

used to assess, with appropriate caution, how securities with varying levels of credit exposure would be priced.

These measures will only become more valuable as more time passes and new CRT issuances broaden the data set. Especially useful would be selling credit risk associated with mortgages that are currently not included in the CRT program or in the single-family data set.[15] Furthermore, researchers with the availability to track the paydown over time of certain loan pools could dynamically build these distributions and study how uncertainty changes over time, where as I am only able to create them at the very beginning of the issuance.

15. The private sector has engaged in similar transactions that are less standardized in terms of security structure as well as the types of underlying collateral. See here for a list of these types of transactions.

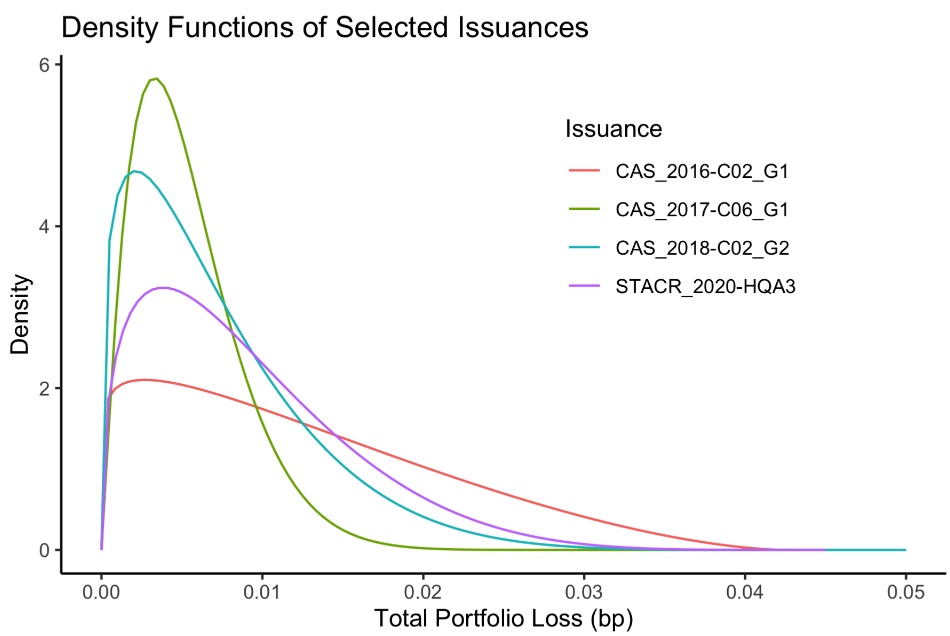



**Figure 1.11:** This plot shows fitted risk-neutral distributions for several of the bond issuances in the sample. This plot shows that tail-risk was heightened in 2016, when credit spreads in markets increased as a result of geopolitical uncertainty.

### 1.5.2 Implied G-fees

This section describes the main implication of calculating the implied-loss distributions, the calculation of a market implied g-fee. Recall from equation 2.12 that the expected losses from the risk-neutral distribution contain both the expected losses and the risk premium component. Thus, in order to capture the implied credit cost portion of the g-fee we take the expected value of the loss-distribution (simply $\alpha/(\alpha+\beta)$ for a beta distribution, and divide by the weighted average life of the mortgage pool, where $i$ denotes the particular bond issuance denoted by a GSE, group, and issuance date:

$$g - fee_i = \frac{\mathbb{E}^{\mathbb{Q}}[Loss_i]}{WAL_i} = \frac{\alpha_i}{[\alpha_i + \beta_i] \times WAL_i} \tag{1.12}$$

Figure 1.12 below shows the result of performing this exercise for all 121 issuances in the data set and is the main result of the paper. I have provided a simple method for extracting an implied loss distribution and g-fee from CRT bond issuances that is independent of a particular econometric model for losses, but rather uses tranche market prices themselves to calculate expected losses.

In figure 1.12, implied g-fees are grouped by GSE and LTV group to be consistent with the summary statistics provided earlier. Intuitively, g-fees co-move strongly over time. From eyeballing the graph, it seems like prices have stabilized over time and now give a pretty consistent view of the implied loss function at least as compared to earlier in the sample when the estimated appear noisier, at least in the cross-section.

Another immediately noticeable feature of the results is that g-fees drop significantly in the beginning of 2017. This coincides with the cessation of selling the very first-loss piece, and the dropping of credit-spreads in the economy at large, so it is difficult to disentangle the two. Lower g-fees in later years, however, are not the product of the first loss-piece not being sold as the standard attachment point for the

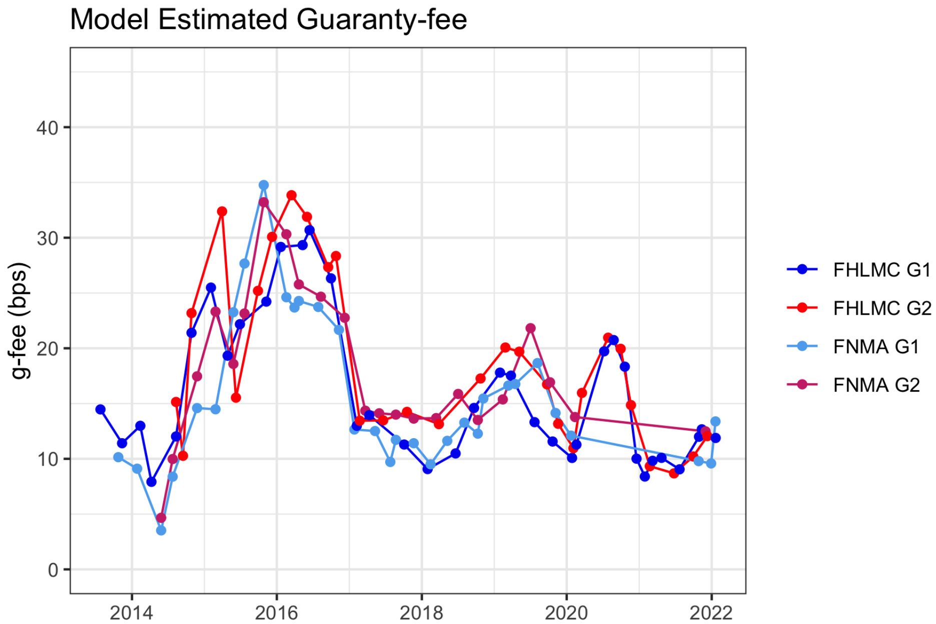



**Figure 1.12:** This plot shows the expected value of the g-fee in the fitted distributions, calculated as the risk-neutral expected losses divided by the weighted average life of the mortgage pool.

B-tranche is now just 15 basis points. Even if we assumed the first 15 basis points was always lost, the g-fee could increase at most 3 basis points at an assumed WAL of 5. This would still keep the implied g-fee generally below the charged levels. In section 1.5.4 below, I will go into further detail about this implied g-fee in relationship to charged g-fee levels and the difference between the two.

### 1.5.3 Correlation with Market Yields

One important outstanding question is the relationship between CRT bond spreads and credit spreads in the wider economy. The existence of common time-varying risk premiums in credit market instruments (Elton et al. (2001), Collin-Dufresne, Goldstein, and Martin (2001)) suggests that one my find a similar result in CRT bonds: that is, since mortgage credit risk is so systematic, we may find that the spreads implied by the model are highly correlated with credit spreads in corporate bonds and elsewhere. To be clear, the implied g-fee measured in this paper is not purely a risk premium component, because it also contains the expectation of losses under the physical measure as well. If I assume that this component is effectively constant over the sample, which is reasonable given the relative stability in housing markets during the period examined, I can compare the implied g-fee to market wide credit spreads by normalizing each to their starting values at the beginning of the CRT data sample.

Figure 1.13 plots the time-series of the estimated g-fee series and both the ICE BBB OAS and AAA-BAA spread, normalized to their values on the date of the first CRT issuance in 2013. These spreads were chosen as to be two representative examples of corporate credit spreads and not because they are necessarily the most correlated. One can see that the correlation here is quite strong and is suggestive that the market views household mortgage credit risk as systemic and highly correlated with market-wide measures of credit risk. Monthly changes of the average implied g-fee have a correlation coefficient of 0.21 ($t = 1.96, N = 79$) with monthly changes in the AAA-BAA credit spread, which is a measure of the yield spread across the universe of investment grade bonds. The correlation of changes in the implied g-fee with monthly changes in the ICE BBB OAS measure are 0.06 and not statistically significant.

In unreported results, I do not find statistically significant correlation between changes in implied CRT g-fees and macroeconomic factors, such as unemployment. Nor do I find support that housing related factors, such as Case-Schiller house price appreciation, delinquency rates, or changes in bank lending standards. While this could be due to a small sample size and short history, it could also be due to the fact that the heavy loading on market wide credit risk is capturing these effects that correlate systematically with market wide credit risk.

These results are quite interesting in the context of mortgage credit models. For

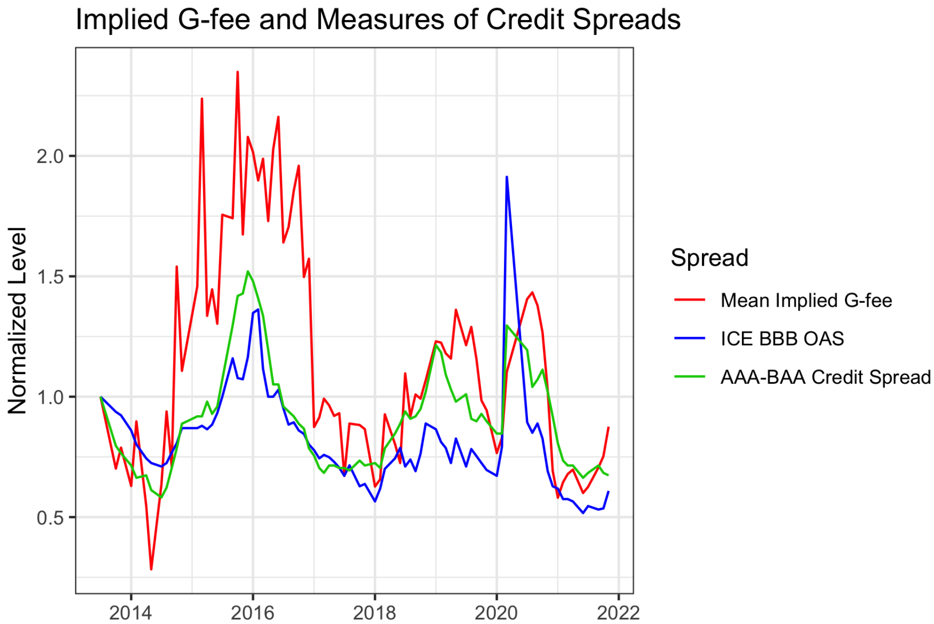



**Figure 1.13:** This plot shows the time-series of implied g-fee (for exposition, averaged at the monthly level when there are multiple bond issuances in particular month)and two measures of credit spreads: the ICE BBB option-adjusted spread index and the AAA-BBB credit spread.

example, Schwartz and Torous (1992) argue in an option theoretic model that the fair g-fee is extremely sensitive to house price volatility. These results suggest those who bear credit risk in mortgage markets may require a significant premium for the systemic component of the risk rather than housing specific factors. More granular data on spreads, however, may be able to discover relationships between these variables and the underlying loans. For example, using third party pricing model data, Gete, Tsouderou, and Wachter (2020) show a relationship between exposure to hurricane Harvey and option-adjusted spreads on CRT tranches. In the following section, I will describe potential explanations for the difference in levels between the implied g-fee and the charged level by the GSEs.

### 1.5.4 Explaining the G-fee Difference

The difference between implied and charged g-fees is shown below in figure 1.14. Unfortunately, as alluded to in the introduction, it is difficult to calculate the exact charged g-fee on a given loan due to the granularity at which the GSEs disclose their realized charged g-fees. Thus the actual charged g-fee on new acquisitions, shown in the red line in the figure, is including loans that may not necessarily be part of CRT collateral. This means that this gap likely serves as an upper bound on the difference between charged and implied levels, as the loans underlying CRT bonds are generally considered to be of higher quality. Several potential explanations come to mind for why there may be a difference between the level of implied g-fees and those charged by the GSEs.

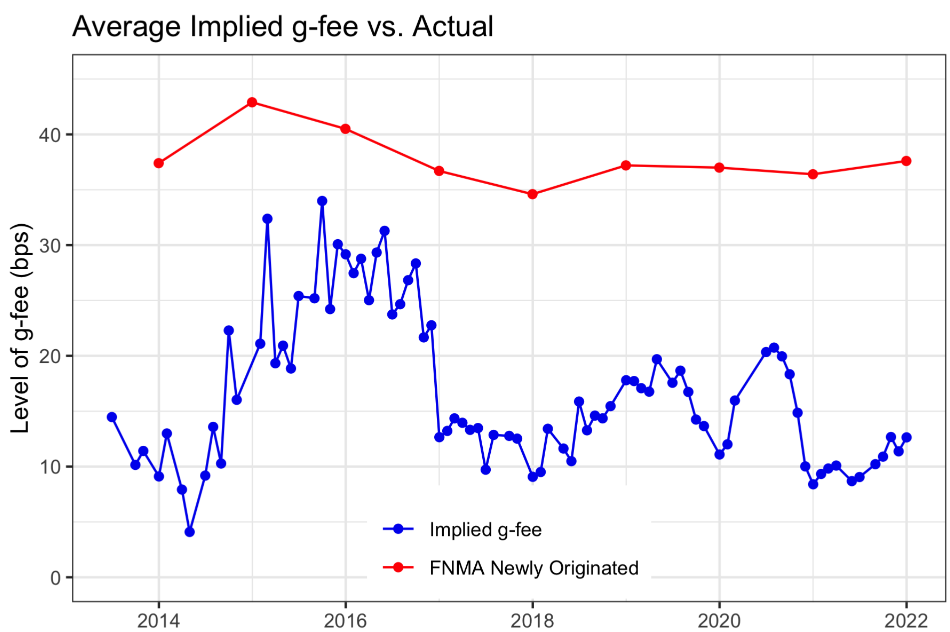



**Figure 1.14:** This plot shows the average expected value of the g-fee in a given month (if there were multiple issuances in that month, regardless of LTV group), plotted alongside the FNMA 10-k reported charged g-fee on newly originated loans. Since the g-fee is reported inclusive of SG&A expenses, I assume they are around 10 basis points in line with Palmer (2017). This results in an implied g-fee that is on average 10-15 basis points lower then the charged level.

One potential explanation is that the implied g-fee is missing compensation for retained tranches that are senior to the subordinated portions sold off in CRT transactions. Unfortunately, a market value for these pieces is impossible to measure. If one were to attribute all of the gap between the charged g-fee and the market implied g-fee as compensation for this risk, it would mean that the market views a fair insurance premium for the super senior portion to be about 10-15 basis points per year. This seems unlikely, especially in the recent CRT issuances, because much of the capital structure is currently covered by CRT bonds; losses above 4% would be extremely unlikely, as they would be higher than losses experienced even during the financial crisis, when underwriting standards were much lower (recall the historical data in section 1.3.1).

Naturally, some variation in the mdoel-based g-fee will be caused by variations in the pieces of risk which are sold. One could imagine the ideal laboratory to test a theory that the difference in levels is due to catastrophic risk would be there were two identical CRT issuances, and one sold off higher detachment points. A deliberately simple way to get at this issue is to compare the shift in Freddie Mac detachment points in 2016 that brought them in line with Fannie Mae issuances. Appendix figure 1.A.8 shows a plot of detachment points, implied g-fees, and a measure of market wide credit spreads over this period. Interestingly, we see that the higher FHLMC g-fee falls to be more inline with the FNMA implied fee after this change. If we take the conservative view that the market required 5-7 basis points for this extra 100 basis points of credit protection, it would simply still not be enough to explain the large gap in implied/actual fee levels in recent years. Upon the availability of more data, including more CRT deals with varying attachment/detachment points, future research could incorporate a study that attempts to exploit variation

Many CRT bonds also contain provisions that in extreme situations, such as high delinquency rates, unscheduled cash flows are no longer returned to subordinate bond holders. Because of the way defaults are modeled in this paper, these provisions are unlikely to have an affect on the results. Since I only consider default scenarios in which entire tranches are written down, the bonds cannot enter the state in which they are limbo between mortgages becoming delinquent and losses being realized.[16] That being said, we cannot strictly rule out that this difference is compensation for the catastrophic event in which the credit coverage of CRT bonds is not adequate.

In a similar vein, another potential explanation is that moral hazard costs are very high. Recall from figure 1.4 that the GSEs retain a vertical slice of a particular tranche (at least 5%, and often times higher on junior bonds). The model described in this paper assumes that the entire tranche is traded in the secondary market. These costs, however, would have to be very large to explain this spread. Although possible, this too seems unlikely because full loan level data is provided on the bonds underlying the deals. This would require the moral hazard to occur in terms of actual fraud on the data or in the underwriting process, and not on the level of choosing which loans were included in the deals.[17]

The most likely explanation is likely the well-known result that g-fees are deliberately high on high quality mortgages. This so-called cross-subsidization of mortgages would occur when the GSEs charge above market rates to insure credit risk on higher quality loans, so that they can pursue mandates related to fair housing and the provision of mortgage credit to lower income/credit quality households (See Cooperstein and Stegman (2019) and Goodman et al. (2022) for empirical discussions of this idea, and Gete and Zecchetto (2017) for a theoretical model). This phenomenon also arises geographically, since the g-fees are not based on loan location, as shown in Hurst et

16. See the O'Neill (2022) appendix on special features and CRT bond prospectuses for more information on these cash flow provisions for special cases.

17. Lai and Van Order (2019) offer a theoretical analysis on managerial incentives for the types of loans pooled into CRT bonds, and Echeverry (2020) empirically analyzes loans sold in credit risk transfer deals in comparison to those retained on GSE balance sheets.

al. (2016). In the next section, I examine the relationship between the model implied g-fees and how they can help quantify the level of cross-subsidization the GSEs are engaging in.

### 1.5.5 Implications for Credit Subsidies

The results of this paper are consistent with the idea that the higher quality borrowers overpay on GSE loans to subsidize higher credit risk borrowers who are likely underpaying for the g-fee on a relative basis. To quantify this gap, I correct the average implied g-fee for potential tail risk concerns as discussed above and in appendix figure 1.A.8 by adding an additional 5 basis points to bond issuances that have subordinate detachment points equal or below 400 basis points. I then take the difference between the average charged g-fee on new acquisitions in a given year and the average implied g-fee from the pricing model. Figure 1.15 below shows the results, which results in a gap of 16 basis points on average. This should be considered an upper bound since average charged g-fees from Fannie Mae's 10k likely include some riskier loans that bring the average rate up.

Intuitively, the plot looks like a mirror image of credit spreads during this period. All else equal, the cross-subsidy generated by high quality loans such as those underlying the CRT bonds is highest when credit spreads are lowest. When credit spreads increase, the "expected return" on holding mortgage credit increases, and so either less credit is available for affordable housing pursuits or the GSEs (and therefore the government and taxpayer while the GSEs are in conservatorship) are taking on more risk during these times. Having a market view of mortgage credit risk allows us to quantify the level of this spread; for example, 16 basis points are freed up on average that could be used to "underprice" lower quality loans that advance a fair-housing agenda.

This spread can also be used by the GSEs to model when g-fees could be getting

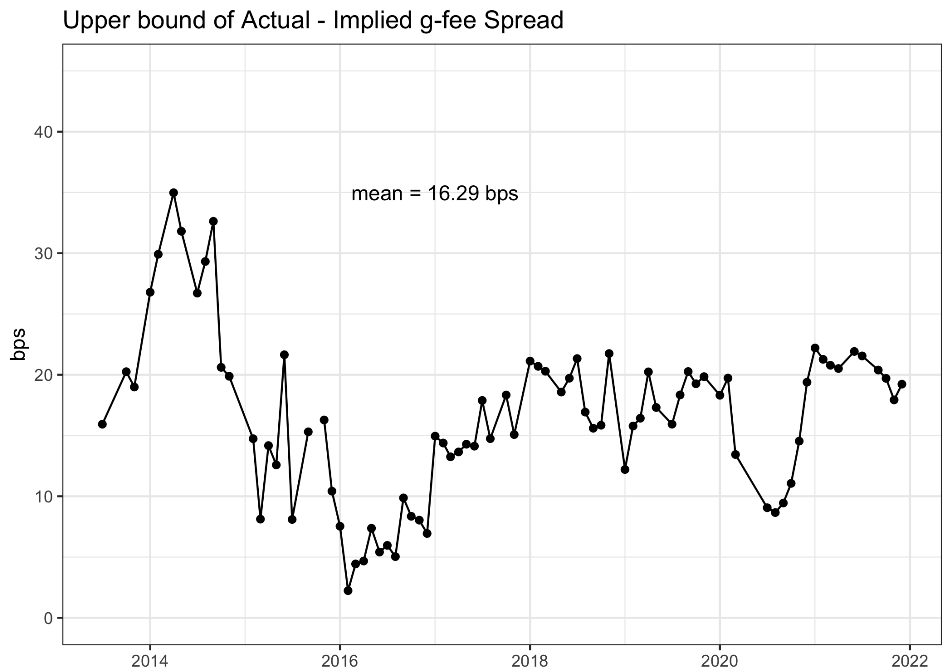



**Figure 1.15:** This figure plots the upper bound of the actual - implied g-fee spread. The implied g-fee is calculated here as the average over CRT bond issuances in a given month, with a 5 basis points adjustment for potential retained risk as described above in section 1.5.4.

less competitive. The GSEs must strike a balance between generating cross-subsidies and keeping g-fees on high quality loans low enough so that there is not an adverse selection problem for which loans originators sell. The results from this paper suggest that when credit risk premiums in the market are low, the g-fee is actual charged g-fee higher and the risk of adverse selection will be higher as well. This may also explain times of relatively higher proportions of loans being held on balance sheet during times of lower risk premiums.

### 1.5.6 Does deal LTV matter for the implied g-fee?

One quick exercise that can be done as a result of the model is whether the market requires a premium for high LTV loans, or those in the so-called Group 2 deals. The average implied g-fee for Group 1 deals is 15.6, where as it is 18.3 bp for Group 2 deals.

This statistically significant difference ($t = 2.7$), although small, provides suggestive evidence that the market does not view primary mortgage insurance as a fail safe protection on loans originated at greater than 80% loan to value ratio. The large difference in empirical loss rates, however, as documented in Goodman et al. (2014), suggests that the market believes PMI is largely effective.

## 1.6 Conclusion

In this paper, I introduced a novel and parsimonious model for estimating implied risk-neutral loss distributions from the market prices of CRT bond issuances. The advantage of my methodology is that, unlike empirical models that value each tranche while incorporating an option adjusted spread, I match CRT tranche prices to an underlying distribution of portfolio losses that makes each tranche market price consistent with one another in a no-arbitrage sense. To my knowledge, this is the first paper to take this approach and provide an answer to the question of what the CRT market says about mortgage credit risk.

My results indicate that the market prices the g-fee on loans backing the bond issuances between 10-30 basis points, and around 15 basis points on average. These market implied g-fees are correlated with yield spreads in corporate bond markets, emphasizing the systemic nature of mortgage credit risk. This range has the interesting implication of falling generally below the current average rate charged on loans purchased by the GSEs.

I discuss qualitatively potential explanations for this gap, including GSE retained tail risk, moral hazard problems, and cross-subsidization across the GSE loan credit stack. I then calculate this gap and find it to be roughly 16 basis points on average, which I argue is a conservative upper bound after correcting for potential concerns about retained risk. This number compares to the level of cross-subsidization coming from high quality loans in Cooperstein and Stegman (2019) and Goodman et al. (2022), but as the added implication that there is variation over time. During times of high risk-premia, the economic value of the subsidy is lower.

There are numerous avenues for future research in this asset class, and the future issuance of new CRT bonds that further span the GSE capital structure will aid in the development of models that can extract their information. Market based indicators of mortgage credit risk are important test cases before a rushed conclusion about the future role of the GSEs in housing finance is made.

# 1.A Appendix

## 1.A.1 Data Sources

| Data | Frequency | Definition.source |
|---|---|---|
| AAA-BAA Credit Spread | Daily | Difference between AAA and BAA credit spreads from FRED (AAA, BAA) |
| Discount Function D(T) | Daily | Using daily Treasury CMT data from FRED, bootstrapped zero-coupon discount curves out to 30 years following the methodology of Longstaff, Mithal, and Neis (2005) |
| FHLMC STACRS Information | Issuance | Information on issuance dates, maturities, CUSIPS, spreads, and more hand-collected Freddie Mac through STACRS deal documents and Freddie Mac Clarity |
| FNMA CAS Information | Issuance | Information on issuance dates, maturities, CUSIPS, spreads, and more hand-collected Fannie Mae through CAS deal documents and Fannie Mae Data Dynamics |
| G-fee History | Yearly | Retreived from yearly Fannie Mae 10K reports |
| ICE BBB Option-adjusted Spread | Daily | Retreived from FRED (BAMLC0A4CBBB) |
| Tranche Market Prices | Daily | TRACE Agency Dataset, WRDS |

**Table 1.4:** This table provides data sources for the paper, along with the frequency at which they are measured and the place at which they can be found.

### 1.A.2 Definitions

| Acronym | Description |
|---|---|
| CAS | Connecticut Avenue Securities. The CRT bond issuance program of Fannie Mae. |
| CIRT | Credit Insurance Risk Transfer. |
| CRT | Credit Risk Transfer. |
| DNA | FHLMC, Group 1 (60-80 LTV), actual-loss deals. |
| FHLMC | Federal Home Loan Mortgage Corporation (Freddie Mac). |
| FNMA | Federal National Mortgage Association (Fannie Mae). |
| GNMA | Government National Mortgage Association Ginnie Mae |
| Group | Refers to the LTV bracket of the underlying loans. Both FNMA and FHLMC issue CRT bonds under this distinction. |
| HQ | The original CRT notes issued by FHLMC. Group 1 (60-80 LTV), fixed-severity deals. |
| HQA | FHLMC, Group 2 (80-97 LTV), actual-loss deals. |
| LIBOR | London Interbank Overnight Rate. The reference rate for the floater spread paid on almost all CRT bonds. |
| REMIC | Real Estate Mortgage Investment Conduit. |
| SOFR | Secured Overnight Financing Rate. The reference rate for CRT coupons since late 2019. |

**Table 1.5:** This table provides a glossary and description of commonly used acronyms. Although I aim to describe acronyms in the paper as they arise, this table aims to provide clarifications on acronyms frequently discussed in other papers, articles, and press releases related to the CRT market.

### 1.A.3 Issuance Summary Statistics by Bond Series

This section presents summary statistics for the bond issuances valued in this paper. To reiterate, the bond issuances presented below are for a particular group of reference loans. The tranching of the particular group of loans varies across the time-series and cross-section of issuances. Summary statistics in these tables indicate, the group and series names, the bond maturities, the tenor of credit protection, the attachment and detachment points of credit protection, and statistics related to the underlying mortgage pool. Table 1.6 shows summary statistics for Freddie Mac STACR Bonds in the Group 1 issuance series. Summary statistics for Freddie Mac STACR Group 2 bonds are presented in table 1.7. Table 1.8 and 1.9 present the Fannie Mae CAS Group 1 and 2 deals.

| Group | Series | Issuance | Maturity | Tenor | Bonds | Attach | Detach | WAC | WLTV | WCS | WAM | WLTI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FHLMC-G1 | STACR 2013-DN01 | 2013-07-26 | 2023-07-25 | 10.0 | 2 | 30 | 165 | 3.82 | 75.4 | 766.8 | 350.0 | 32.2 |
| FHLMC-G1 | STACR 2013-DN02 | 2013-11-12 | 2023-11-25 | 10.0 | 2 | 30 | 195 | 3.57 | 74.9 | 765.3 | 353.0 | 32.1 |
| FHLMC-G1 | STACR 2014-DN01 | 2014-02-12 | 2024-02-25 | 10.0 | 3 | 30 | 350 | 3.69 | 75.3 | 763.1 | 353.0 | 32.6 |
| FHLMC-G1 | STACR 2014-DN02 | 2014-04-09 | 2024-04-25 | 10.0 | 3 | 30 | 350 | 3.89 | 75.7 | 761.3 | 354.0 | 32.8 |
| FHLMC-G1 | STACR 2014-DN03 | 2014-08-11 | 2024-08-25 | 10.0 | 3 | 40 | 360 | 4.53 | 76.1 | 751.8 | 353.0 | 34.6 |
| FHLMC-G1 | STACR 2014-DN04 | 2014-10-28 | 2024-10-25 | 10.0 | 3 | 50 | 420 | 4.58 | 75.9 | 749.2 | 353.0 | 35.0 |
| FHLMC-G1 | STACR 2015-DN01 | 2015-02-03 | 2025-01-25 | 10.0 | 4 | 0 | 350 | 4.53 | 76.3 | 751.3 | 353.0 | 34.6 |
| FHLMC-G1 | STACR 2015-DNA1 | 2015-04-28 | 2027-10-25 | 12.5 | 4 | 0 | 325 | 3.65 | 75.2 | 766.5 | 352.5 | 32.1 |
| FHLMC-G1 | STACR 2015-DNA2 | 2015-06-29 | 2027-12-25 | 12.5 | 4 | 0 | 450 | 4.39 | 76.3 | 750.5 | 352.0 | 34.7 |
| FHLMC-G1 | STACR 2015-DNA3 | 2015-11-09 | 2028-04-25 | 12.5 | 4 | 0 | 485 | 4.11 | 75.6 | 753.2 | 352.0 | 34.6 |
| FHLMC-G1 | STACR 2016-DNA1 | 2016-01-21 | 2028-07-25 | 12.5 | 4 | 0 | 395 | 3.96 | 75.7 | 755.6 | 354.0 | 34.4 |
| FHLMC-G1 | STACR 2016-DNA2 | 2016-05-10 | 2028-10-25 | 12.5 | 4 | 0 | 415 | 4.16 | 76.1 | 751.4 | 353.0 | 35.1 |
| FHLMC-G1 | STACR 2016-DNA3 | 2016-06-14 | 2028-12-25 | 12.5 | 4 | 0 | 400 | 4.20 | 75.9 | 747.4 | 355.0 | 35.2 |
| FHLMC-G1 | STACR 2016-DNA4 | 2016-09-30 | 2029-03-25 | 12.5 | 4 | 0 | 400 | 4.18 | 75.6 | 747.3 | 354.0 | 35.5 |
| FHLMC-G1 | STACR 2017-DNA1 | 2017-02-07 | 2029-07-25 | 12.5 | 4 | 0 | 255 | 3.97 | 75.7 | 750.4 | 352.0 | 35.0 |
| FHLMC-G1 | STACR 2017-DNA2 | 2017-04-11 | 2029-10-25 | 12.5 | 4 | 0 | 230 | 3.82 | 75.7 | 751.6 | 354.0 | 34.7 |
| FHLMC-G1 | STACR 2017-DNA3 | 2017-10-04 | 2030-03-25 | 12.5 | 3 | 50 | 250 | 3.94 | 75.3 | 749.4 | 352.0 | 35.3 |
| FHLMC-G1 | STACR 2018-DNA1 | 2018-01-30 | 2030-07-25 | 12.5 | 3 | 50 | 310 | 4.43 | 76.1 | 742.7 | 353.0 | 36.2 |
| FHLMC-G1 | STACR 2018-DNA2 | 2018-06-20 | 2030-12-25 | 12.5 | 3 | 50 | 250 | 4.22 | 76.1 | 745.7 | 352.0 | 35.9 |
| FHLMC-G1 | STACR 2018-DNA3 | 2018-09-21 | 2048-09-25 | 30.0 | 4 | 10 | 300 | 4.34 | 75.8 | 742.0 | 354.0 | 36.5 |
| FHLMC-G1 | STACR 2019-DNA1 | 2019-01-30 | 2049-01-25 | 30.0 | 4 | 10 | 300 | 4.76 | 76.0 | 739.6 | 353.0 | 37.1 |
| FHLMC-G1 | STACR 2019-DNA2 | 2019-03-26 | 2049-03-25 | 30.0 | 4 | 10 | 350 | 4.92 | 76.2 | 740.1 | 354.0 | 37.1 |
| FHLMC-G1 | STACR 2019-DNA3 | 2019-07-23 | 2049-07-25 | 30.0 | 4 | 10 | 325 | 4.96 | 76.0 | 741.3 | 351.0 | 36.7 |
| FHLMC-G1 | STACR 2019-DNA4 | 2019-10-22 | 2049-10-25 | 30.0 | 4 | 10 | 300 | 4.98 | 75.9 | 740.1 | 350.0 | 36.5 |
| FHLMC-G1 | STACR 2020-DNA1 | 2020-01-28 | 2050-01-25 | 30.0 | 4 | 10 | 275 | 4.60 | 76.0 | 745.0 | 350.0 | 36.3 |
| FHLMC-G1 | STACR 2020-DNA2 | 2020-02-19 | 2050-02-25 | 30.0 | 4 | 10 | 250 | 4.19 | 76.1 | 750.8 | 352.0 | 35.7 |
| FHLMC-G1 | STACR 2020-DNA3 | 2020-07-08 | 2050-06-25 | 30.0 | 4 | 25 | 300 | 3.94 | 75.5 | 753.3 | 352.0 | 35.3 |
| FHLMC-G1 | STACR 2020-DNA4 | 2020-08-25 | 2050-08-25 | 30.0 | 4 | 25 | 300 | 3.95 | 75.5 | 752.4 | 353.0 | 35.7 |
| FHLMC-G1 | STACR 2020-DNA5 | 2020-10-20 | 2050-10-25 | 30.0 | 4 | 10 | 250 | 3.57 | 74.8 | 758.1 | 354.0 | 33.8 |
| FHLMC-G1 | STACR 2020-DNA6 | 2020-12-18 | 2050-12-25 | 30.0 | 4 | 25 | 200 | 3.40 | 74.4 | 758.2 | 354.0 | 33.3 |
| FHLMC-G1 | STACR 2021-DNA1 | 2021-01-29 | 2051-01-25 | 30.0 | 4 | 25 | 200 | 3.26 | 74.5 | 762.1 | 353.0 | 33.1 |
| FHLMC-G1 | STACR 2021-DNA2 | 2021-03-09 | 2033-08-25 | 12.5 | 4 | 25 | 200 | 3.09 | 74.3 | 762.5 | 354.0 | 33.3 |
| FHLMC-G1 | STACR 2021-DNA3 | 2021-04-23 | 2033-10-25 | 12.5 | 4 | 25 | 200 | 2.97 | 74.2 | 763.4 | 353.0 | 33.1 |
| FHLMC-G1 | STACR 2021-DNA5 | 2021-07-23 | 2034-01-25 | 12.5 | 4 | 25 | 175 | 2.94 | 73.9 | 762.8 | 352.0 | 33.1 |
| FHLMC-G1 | STACR 2021-DNA6 | 2021-10-29 | 2041-10-25 | 20.0 | 4 | 25 | 175 | 2.86 | 73.9 | 759.9 | 351.0 | 33.5 |
| FHLMC-G1 | STACR 2021-DNA7 | 2021-11-12 | 2041-11-25 | 20.0 | 4 | 25 | 175 | 3.03 | 74.3 | 755.2 | 354.0 | 33.9 |
| FHLMC-G1 | STACR 2022-DNA1 | 2022-01-21 | 2042-01-25 | 20.0 | 5 | 25 | 300 | 3.15 | 74.9 | 752.0 | 355.0 | 34.5 |

**Table 1.6:** Issuance Summary Statistics: FHLMC Group 1

| Group | Series | Issuance | Maturity | Tenor | Bonds | Attach | Detach | WAC | WLTV | WCS | WAM | WLTI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FNMA-G1 | CAS 2013-C01 | 2013-10-24 | 2023-10-25 | 10.0 | 2 | 30 | 165 | 3.85 | 74.9 | 770.5 | 347.9 | 31.8 |
| FNMA-G1 | CAS 2014-C01 | 2014-01-27 | 2024-01-25 | 10.0 | 2 | 30 | 165 | 3.64 | 74.5 | 769.8 | 347.9 | 31.5 |
| FNMA-G1 | CAS 2014-C02 | 2014-05-28 | 2024-05-25 | 10.0 | 2 | 30 | 175 | 3.59 | 74.3 | 768.7 | 346.6 | 31.7 |
| FNMA-G1 | CAS 2014-C03 | 2014-07-25 | 2024-07-25 | 10.0 | 2 | 30 | 200 | 3.75 | 74.8 | 766.0 | 347.8 | 32.5 |
| FNMA-G1 | CAS 2014-C04 | 2014-11-25 | 2024-11-25 | 10.0 | 2 | 30 | 200 | 4.01 | 75.3 | 762.6 | 346.2 | 33.2 |
| FNMA-G1 | CAS 2015-C01 | 2015-02-26 | 2025-02-25 | 10.0 | 2 | 40 | 215 | 4.56 | 75.9 | 758.0 | 345.6 | 34.2 |
| FNMA-G1 | CAS 2015-C02 | 2015-05-27 | 2025-05-25 | 10.0 | 2 | 40 | 275 | 4.60 | 76.0 | 752.6 | 346.6 | 34.7 |
| FNMA-G1 | CAS 2015-C03 | 2015-07-22 | 2025-07-25 | 10.0 | 2 | 40 | 280 | 4.51 | 76.2 | 752.7 | 349.5 | 34.3 |
| FNMA-G1 | CAS 2015-C04 | 2015-10-27 | 2028-04-25 | 12.5 | 2 | 50 | 305 | 4.42 | 76.0 | 751.7 | 349.8 | 34.5 |
| FNMA-G1 | CAS 2016-C01 | 2016-02-18 | 2028-08-25 | 12.5 | 3 | 0 | 285 | 4.18 | 75.2 | 754.0 | 349.2 | 34.2 |
| FNMA-G1 | CAS 2016-C02 | 2016-03-30 | 2028-09-25 | 12.5 | 3 | 0 | 275 | 4.00 | 74.9 | 757.6 | 350.8 | 33.8 |
| FNMA-G1 | CAS 2016-C03 | 2016-04-21 | 2028-10-25 | 12.5 | 3 | 0 | 260 | 4.03 | 75.4 | 756.2 | 351.9 | 33.8 |
| FNMA-G1 | CAS 2016-C04 | 2016-07-28 | 2029-01-25 | 12.5 | 3 | 0 | 275 | 4.21 | 75.7 | 753.5 | 351.3 | 34.1 |
| FNMA-G1 | CAS 2016-C06 | 2016-11-09 | 2029-04-25 | 12.5 | 3 | 0 | 275 | 4.20 | 75.4 | 751.5 | 352.2 | 34.3 |
| FNMA-G1 | CAS 2017-C01 | 2017-01-26 | 2029-07-25 | 12.5 | 3 | 50 | 265 | 4.02 | 75.2 | 754.3 | 353.4 | 34.1 |
| FNMA-G1 | CAS 2017-C03 | 2017-05-10 | 2029-10-25 | 12.5 | 3 | 50 | 255 | 3.82 | 75.0 | 757.8 | 353.2 | 33.5 |
| FNMA-G1 | CAS 2017-C05 | 2017-07-26 | 2030-01-25 | 12.5 | 3 | 50 | 290 | 3.76 | 74.7 | 757.1 | 352.9 | 33.6 |
| FNMA-G1 | CAS 2017-C06 | 2017-08-23 | 2030-02-25 | 12.5 | 3 | 50 | 280 | 4.08 | 74.9 | 751.4 | 354.1 | 34.6 |
| FNMA-G1 | CAS 2017-C07 | 2017-11-21 | 2030-05-25 | 12.5 | 3 | 50 | 305 | 4.38 | 75.3 | 750.1 | 353.4 | 35.1 |
| FNMA-G1 | CAS 2018-C01 | 2018-02-14 | 2030-07-25 | 12.4 | 3 | 50 | 310 | 4.34 | 75.8 | 752.1 | 355.0 | 34.9 |
| FNMA-G1 | CAS 2018-C03 | 2018-05-09 | 2030-10-25 | 12.5 | 3 | 50 | 320 | 4.25 | 75.5 | 748.5 | 355.1 | 36.0 |
| FNMA-G1 | CAS 2018-C05 | 2018-08-03 | 2031-01-25 | 12.5 | 3 | 50 | 335 | 4.28 | 75.3 | 748.6 | 355.1 | 36.9 |
| FNMA-G1 | CAS 2018-C06 | 2018-10-10 | 2031-03-25 | 12.5 | 3 | 50 | 370 | 4.60 | 75.4 | 746.6 | 355.9 | 37.4 |
| FNMA-G1 | CAS 2018-R07 | 2018-11-07 | 2031-04-25 | 12.5 | 3 | 50 | 385 | 4.80 | 75.7 | 747.8 | 356.9 | 37.5 |
| FNMA-G1 | CAS 2019-R02 | 2019-03-13 | 2031-08-25 | 12.5 | 3 | 50 | 370 | 4.89 | 75.8 | 750.3 | 355.1 | 37.3 |
| FNMA-G1 | CAS 2019-R03 | 2019-04-17 | 2031-09-25 | 12.4 | 3 | 50 | 370 | 4.95 | 75.8 | 749.1 | 356.5 | 37.6 |
| FNMA-G1 | CAS 2019-R05 | 2019-08-07 | 2039-07-25 | 20.0 | 3 | 25 | 365 | 5.07 | 75.6 | 741.0 | 356.0 | 38.0 |
| FNMA-G1 | CAS 2019-R07 | 2019-11-06 | 2039-10-25 | 20.0 | 3 | 25 | 325 | 4.61 | 75.5 | 747.0 | 356.0 | 37.0 |
| FNMA-G1 | CAS 2020-R01 | 2020-01-23 | 2040-01-25 | 20.0 | 3 | 20 | 285 | 4.28 | 75.4 | 755.6 | 355.7 | 36.0 |
| FNMA-G1 | CAS 2021-R01 | 2021-10-27 | 2041-10-25 | 20.0 | 4 | 25 | 160 | 2.98 | 74.0 | 761.0 | 350.0 | 34.0 |
| FNMA-G1 | CAS 2021-R03 | 2021-12-29 | 2041-12-25 | 20.0 | 4 | 25 | 220 | 2.91 | 73.8 | 760.0 | 350.0 | 34.0 |
| FNMA-G1 | CAS 2022-R01 | 2022-01-20 | 2041-12-25 | 19.9 | 4 | 25 | 230 | 2.88 | 73.6 | 758.0 | 351.0 | 34.0 |

**Table 1.7:** Issuance Summary Statistics: FNMA Group 1

| Group | Series | Issuance | Maturity | Tenor | Bonds | Attach | Detach | WAC | WLTV | WCS | WAM | WLTI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FHLMC-G2 | STACR 2014-HQ01 | 2014-08-11 | 2024-08-25 | 10.0 | 3 | 75 | 410 | 4.53 | 92.1 | 750.4 | 353.0 | 34.6 |
| FHLMC-G2 | STACR 2014-HQ02 | 2014-09-15 | 2024-09-25 | 10.0 | 3 | 60 | 410 | 3.71 | 91.4 | 758.6 | 346.0 | 32.8 |
| FHLMC-G2 | STACR 2014-HQ03 | 2014-10-28 | 2024-10-25 | 10.0 | 3 | 85 | 475 | 4.56 | 92.2 | 746.6 | 353.0 | 34.7 |
| FHLMC-G2 | STACR 2015-HQ01 | 2015-03-31 | 2025-03-25 | 10.0 | 4 | 0 | 475 | 4.53 | 92.4 | 746.8 | 352.0 | 34.5 |
| FHLMC-G2 | STACR 2015-HQ02 | 2015-06-09 | 2025-05-25 | 10.0 | 4 | 0 | 410 | 3.71 | 91.4 | 758.9 | 337.0 | 32.8 |
| FHLMC-G2 | STACR 2015-HQA1 | 2015-09-28 | 2028-03-25 | 12.5 | 4 | 0 | 495 | 4.40 | 92.3 | 745.0 | 349.0 | 34.8 |
| FHLMC-G2 | STACR 2015-HQA2 | 2015-12-08 | 2028-05-25 | 12.5 | 4 | 0 | 540 | 4.16 | 91.7 | 747.0 | 351.0 | 35.0 |
| FHLMC-G2 | STACR 2016-HQA1 | 2016-03-15 | 2028-09-25 | 12.5 | 4 | 0 | 440 | 3.98 | 91.7 | 749.4 | 352.0 | 34.5 |
| FHLMC-G2 | STACR 2016-HQA2 | 2016-06-01 | 2028-11-25 | 12.5 | 4 | 0 | 450 | 4.16 | 92.0 | 747.7 | 352.0 | 34.9 |
| FHLMC-G2 | STACR 2016-HQA3 | 2016-09-16 | 2029-03-25 | 12.5 | 4 | 0 | 420 | 4.16 | 91.9 | 745.8 | 352.0 | 35.1 |
| FHLMC-G2 | STACR 2016-HQA4 | 2016-10-25 | 2029-04-25 | 12.5 | 4 | 0 | 428 | 4.15 | 91.9 | 743.6 | 354.0 | 35.7 |
| FHLMC-G2 | STACR 2017-HQA1 | 2017-02-22 | 2029-08-25 | 12.5 | 4 | 0 | 325 | 3.92 | 92.0 | 746.3 | 353.0 | 35.3 |
| FHLMC-G2 | STACR 2017-HQA2 | 2017-06-20 | 2029-12-25 | 12.5 | 3 | 50 | 300 | 3.76 | 91.8 | 746.2 | 353.0 | 35.6 |
| FHLMC-G2 | STACR 2017-HQA3 | 2017-10-18 | 2030-04-25 | 12.5 | 3 | 50 | 370 | 3.97 | 91.7 | 745.1 | 353.0 | 36.0 |
| FHLMC-G2 | STACR 2018-HQA1 | 2018-03-28 | 2030-09-25 | 12.5 | 3 | 50 | 320 | 4.30 | 92.6 | 741.6 | 353.0 | 36.5 |
| FHLMC-G2 | STACR 2018-HQA2 | 2018-10-24 | 2048-10-25 | 30.0 | 4 | 10 | 300 | 4.19 | 92.6 | 741.0 | 351.0 | 37.0 |
| FHLMC-G2 | STACR 2019-HQA1 | 2019-02-26 | 2049-02-25 | 30.0 | 4 | 10 | 360 | 4.67 | 92.9 | 737.5 | 353.0 | 37.6 |
| FHLMC-G2 | STACR 2019-HQA2 | 2019-05-07 | 2049-04-25 | 30.0 | 4 | 10 | 350 | 4.82 | 93.0 | 736.9 | 352.0 | 38.0 |
| FHLMC-G2 | STACR 2019-HQA3 | 2019-09-24 | 2049-09-25 | 30.0 | 4 | 10 | 350 | 4.92 | 92.9 | 736.7 | 351.0 | 38.1 |
| FHLMC-G2 | STACR 2019-HQA4 | 2019-11-19 | 2049-11-25 | 30.0 | 4 | 10 | 325 | 4.98 | 92.7 | 736.7 | 352.0 | 38.3 |
| FHLMC-G2 | STACR 2020-HQA1 | 2020-02-04 | 2050-01-25 | 30.0 | 4 | 10 | 300 | 4.51 | 92.8 | 741.9 | 354.0 | 37.4 |
| FHLMC-G2 | STACR 2020-HQA2 | 2020-03-18 | 2050-03-25 | 30.0 | 4 | 10 | 300 | 4.12 | 92.6 | 745.9 | 354.0 | 36.8 |
| FHLMC-G2 | STACR 2020-HQA3 | 2020-07-28 | 2050-07-25 | 30.0 | 4 | 25 | 330 | 3.85 | 92.0 | 750.8 | 352.0 | 36.1 |
| FHLMC-G2 | STACR 2020-HQA4 | 2020-09-29 | 2050-09-25 | 30.0 | 4 | 25 | 300 | 3.86 | 92.0 | 750.4 | 353.0 | 36.3 |
| FHLMC-G2 | STACR 2020-HQA5 | 2020-11-20 | 2050-11-25 | 30.0 | 4 | 25 | 275 | 3.45 | 91.2 | 752.8 | 355.0 | 34.8 |
| FHLMC-G2 | STACR 2021-HQA1 | 2021-02-23 | 2033-08-25 | 12.5 | 4 | 25 | 250 | 3.16 | 90.9 | 755.1 | 355.0 | 34.8 |
| FHLMC-G2 | STACR 2021-HQA2 | 2021-06-25 | 2033-12-25 | 12.5 | 4 | 25 | 250 | 2.92 | 90.7 | 755.9 | 353.0 | 34.7 |
| FHLMC-G2 | STACR 2021-HQA3 | 2021-09-30 | 2041-09-25 | 20.0 | 4 | 25 | 200 | 2.81 | 90.4 | 753.8 | 353.0 | 34.8 |
| FHLMC-G2 | STACR 2021-HQA4 | 2021-12-10 | 2041-12-25 | 20.0 | 4 | 25 | 200 | 3.00 | 91.1 | 752.5 | 354.0 | 35.0 |

**Table 1.8:** Issuance Summary Statistics: FHLMC Group 2

| Group | Series | Issuance | Maturity | Tenor | Bonds | Attach | Detach | WAC | WLTV | WCS | WAM | WLTI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| FNMA-G2 | CAS 2014-C02 | 2014-05-28 | 2024-05-25 | 10.0 | 2 | 65 | 240 | 3.58 | 91.0 | 761.8 | 346.6 | 32.9 |
| FNMA-G2 | CAS 2014-C03 | 2014-07-25 | 2024-07-25 | 10.0 | 2 | 65 | 240 | 3.74 | 91.5 | 760.1 | 347.9 | 33.4 |
| FNMA-G2 | CAS 2014-C04 | 2014-11-25 | 2024-11-25 | 10.0 | 2 | 65 | 255 | 4.05 | 92.2 | 758.6 | 346.3 | 34.0 |
| FNMA-G2 | CAS 2015-C01 | 2015-02-26 | 2025-02-25 | 10.0 | 2 | 70 | 280 | 4.61 | 92.6 | 755.2 | 345.7 | 34.9 |
| FNMA-G2 | CAS 2015-C02 | 2015-05-27 | 2025-05-25 | 10.0 | 2 | 80 | 285 | 4.61 | 92.5 | 750.8 | 346.7 | 35.2 |
| FNMA-G2 | CAS 2015-C03 | 2015-07-22 | 2025-07-25 | 10.0 | 2 | 80 | 290 | 4.53 | 92.2 | 750.0 | 349.5 | 34.9 |
| FNMA-G2 | CAS 2015-C04 | 2015-10-27 | 2028-04-25 | 12.5 | 2 | 80 | 310 | 4.44 | 92.0 | 749.6 | 349.7 | 35.0 |
| FNMA-G2 | CAS 2016-C01 | 2016-02-18 | 2028-08-25 | 12.5 | 2 | 100 | 290 | 4.25 | 91.6 | 750.0 | 349.1 | 34.9 |
| FNMA-G2 | CAS 2016-C03 | 2016-04-21 | 2028-10-25 | 12.5 | 3 | 0 | 300 | 4.06 | 91.8 | 752.4 | 350.5 | 34.4 |
| FNMA-G2 | CAS 2016-C05 | 2016-08-10 | 2029-01-25 | 12.5 | 3 | 0 | 295 | 4.21 | 92.2 | 751.3 | 352.1 | 34.7 |
| FNMA-G2 | CAS 2016-C07 | 2016-12-08 | 2029-05-25 | 12.5 | 3 | 0 | 310 | 4.13 | 92.0 | 750.0 | 353.4 | 34.9 |
| FNMA-G2 | CAS 2017-C02 | 2017-03-22 | 2029-09-25 | 12.5 | 3 | 50 | 300 | 3.87 | 92.1 | 752.1 | 353.6 | 34.6 |
| FNMA-G2 | CAS 2017-C04 | 2017-05-31 | 2029-11-25 | 12.5 | 3 | 50 | 310 | 3.73 | 91.8 | 752.4 | 355.2 | 34.7 |
| FNMA-G2 | CAS 2017-C06 | 2017-08-23 | 2030-02-25 | 12.5 | 3 | 50 | 345 | 4.26 | 92.2 | 749.1 | 355.8 | 35.6 |
| FNMA-G2 | CAS 2017-C07 | 2017-11-21 | 2030-05-25 | 12.5 | 3 | 50 | 340 | 4.39 | 92.5 | 749.6 | 355.6 | 35.5 |
| FNMA-G2 | CAS 2018-C02 | 2018-03-14 | 2030-08-25 | 12.5 | 3 | 50 | 375 | 4.26 | 92.5 | 749.2 | 355.5 | 35.8 |
| FNMA-G2 | CAS 2018-C04 | 2018-07-03 | 2030-12-25 | 12.5 | 3 | 50 | 370 | 4.21 | 92.5 | 747.6 | 355.1 | 37.6 |
| FNMA-G2 | CAS 2018-C06 | 2018-10-10 | 2031-03-25 | 12.5 | 3 | 50 | 350 | 4.34 | 92.5 | 748.2 | 355.0 | 38.2 |
| FNMA-G2 | CAS 2019-R01 | 2019-02-13 | 2031-07-25 | 12.5 | 3 | 50 | 340 | 4.78 | 92.8 | 749.0 | 355.2 | 38.3 |
| FNMA-G2 | CAS 2019-R04 | 2019-07-03 | 2039-06-25 | 20.0 | 3 | 25 | 365 | 4.89 | 92.8 | 747.7 | 353.8 | 38.5 |
| FNMA-G2 | CAS 2019-R06 | 2019-10-09 | 2039-09-25 | 20.0 | 3 | 25 | 365 | 4.67 | 92.8 | 742.0 | 356.0 | 38.0 |
| FNMA-G2 | CAS 2020-R02 | 2020-02-12 | 2040-01-25 | 20.0 | 3 | 25 | 335 | 4.18 | 92.3 | 749.3 | 356.5 | 37.6 |
| FNMA-G2 | CAS 2021-R02 | 2021-12-01 | 2041-11-25 | 20.0 | 4 | 35 | 245 | 2.94 | 91.6 | 749.0 | 350.0 | 36.0 |

**Table 1.9:** Issuance Summary Statistics: FNMA Group 2

### 1.A.4 Simplified Example of a CRT Bond

This section shows the example paydowns of three CRT tranches analogous to the issuance discussed in section 1.2.3.1. In this example, the CRT issuance references a hypothetical pool of mortgage bonds worth \$1000, all with a mortgage rate of 4%. Three tranches of CRT bonds are sold in this issuance: 2 mezzanine tranches $M1$ and $M2$, and a first-loss piece $B$. For simplicity, we assume that the loans in the reference pool are backed by the CRT issuance immediately, and therefore make their first payment while already having their credit risk backed by the CRT investors.

Figure 1.16 below shows principal paydowns and loss write-downs in three loss scenarios, which are denoted in each column. The three rows represent tranche balances, cumulative principal return, and total cumulative loss write-downs. All balances are normalized to represent \$100 of bond notional. The underlying can be considered to represent \$1 billion of 30 year mortgages with a 4% interest rate.

In column 1, a catastrophic default event wipes out both the junior bond entirely and 90% of the M2 bond at around the 30th month of mortgage amortization. This significantly reduces the subordinated balance, and therefore principal is returned more slowly to the M1 bond which is still fully intact. Once the M1 bond has been entirely paid off, principal begins to flow to the M2 bond in the amount that is still outstanding. Columns 2 and 3 show a similar event at different times and with different portions of the M2 bond surviving the event.

The key takeaway from these plots is to show how the rate of principal return can vary as a function of cumulative loss write-downs, and how the timing of defaults can affect the paydown of the bonds. In the model I will only consider the possibility that tranches are written down entirely upon a default event. O'Neill (2022) goes into further detail on dynamic default behavior by specifying a two factor no-arbitrage model for default rates.

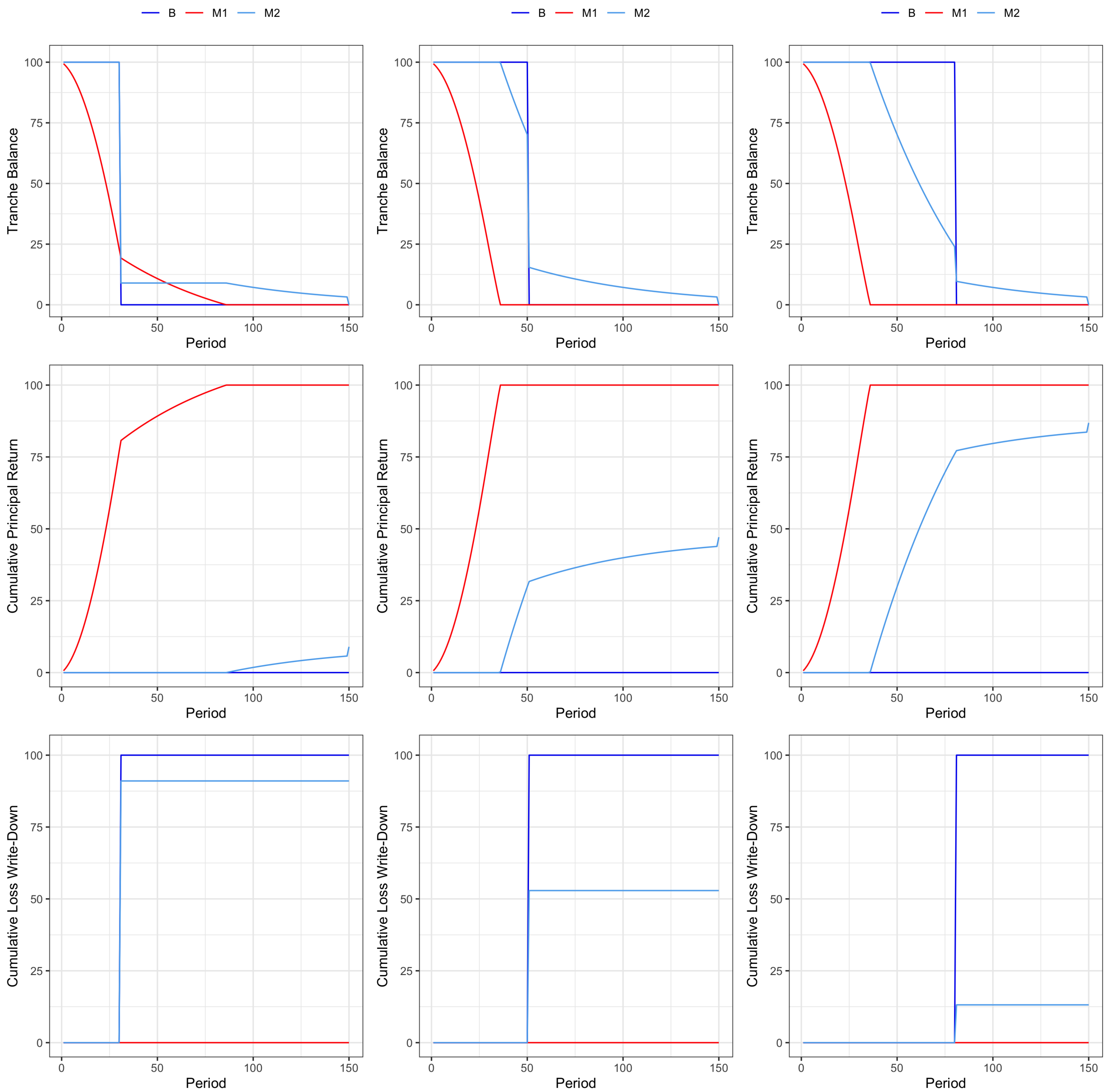



**Figure 1.16:** Examples of CRT Bond Paydowns. Columns represent three different default scenarios and rows represent tranche balances, cumulative principal return, and cumulative loss write-downs in the three default scenarios.

### 1.A.5 Estimation Results

This section shows the results of the estimation exercise for the 121 CRT bond issuances along with summary statistics broken down into each deal LTV group. The RMSE can be interpreted as how well the beta distribution fits the implied probability density function implied by the tranche market prices. A RMSE of 27 basis points means that, for example, an implied default probability of 50% is fit to the beta distribution to the accuracy of 50 ± 0.27%. $\alpha$ and $\beta$ are the implied beta distribution parameters. Also included are the expected weighted average life of the collateral pools under the default model, as well as the total expected losses (risk-neutral) and g-fee, along with their associated distributions.

| | All | Group 1 | Group 2 |
|---|---|---|---|
| N | 121.00 | 69.00 | 52.00 |
| $CPR$ | 0.15 | 0.15 | 0.15 |
| Mean $\mathbb{E}[WAL]$ | 6.16 | 6.17 | 6.14 |
| RMSE (bps) | 27.48 | 23.50 | 32.75 |
| Mean $\alpha$ | 0.90 | 0.87 | 0.95 |
| Mean $\beta$ | 3.62 | 3.48 | 3.81 |
| Mean $\mathbb{E}[Loss]$ | 103.25 | 96.39 | 112.35 |
| SD $\mathbb{E}[Loss]$ | 42.34 | 40.49 | 43.40 |
| Min $\mathbb{E}[Loss]$ | 21.66 | 21.66 | 28.43 |
| Max $\mathbb{E}[Loss]$ | 214.22 | 214.22 | 205.88 |
| Mean $\mathbb{E}[g-fee]$ | 16.77 | 15.63 | 18.29 |
| SD $\mathbb{E}[g-fee]$ | 6.91 | 6.59 | 7.10 |
| Min $\mathbb{E}[g-fee]$ | 3.53 | 3.53 | 4.66 |
| Max $\mathbb{E}[g-fee]$ | 34.77 | 34.77 | 33.85 |

**Table 1.10:** Estimation Results from fitting the beta distribution to the default probabilities implied by the full sample of CRT bond issuances

### 1.A.6 Results from Ignoring the Beta Distribution

This section shows that the results are not heavily dependent on fitting the Beta distribution to the probabilities implied by the pricing model. The plot below shows the g-fee when we directly multiply probabilities by tranche losses. This means that rather than assuming losses meet or exceed a tranche detachment point for a given default state, I assume losses are discrete and are exactly equal to the tranche detachment point at the default time and with the probability given by the model. When the results here are compared to the beta distribution results in 1.12, one can see that fitting the beta distribution cleans up the disparity cross-sectionally, particularly in the earlier part of the sample. Since FNMA did not sell off as much of the credit stack, which understated FHLMC issuances at the same time which had much higher final detachment points. This emphasizes that this model can play a role in reducing the uncertainty of fitting implied loss distributions when the exact portions of credit risk sold may vary over time.

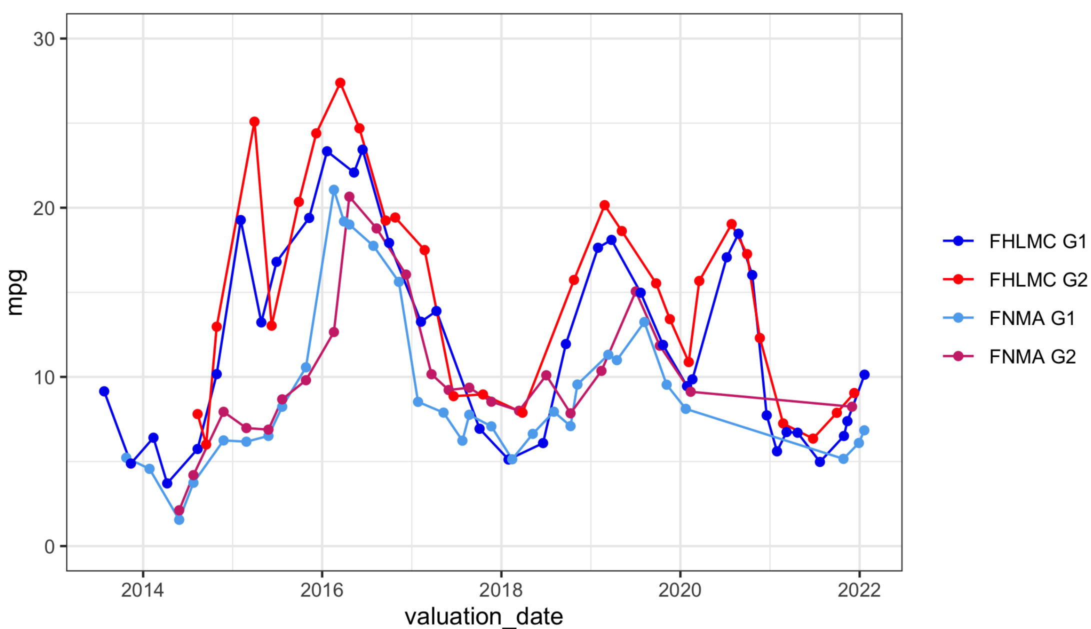



**Figure 1.17:** G-fee Directly From Probabilities

### 1.A.7 Robustness to Prepayment Speeds

This section shows that choice of prepayment rate actually does not have much of an impact on the implied g-fee results. This is because lower (higher) prepayments increase (decrease) the present value of losses, but they also increase the denominator in the g-fee term, the weighted average life of the mortgage pool. The tables below repeat the summary statistics of table 1.10 for alternate prepayment speeds. Panels A-D of the figure show that the implied g-fee for prepayment speeds at 12, 18, 21, and 24% CPR. Recall that the baseline results are calculated using a CPR of 15%.

| | All | Group 1 | Group 2 |
|---|---|---|---|
| N | 121.00 | 69.00 | 52.00 |
| $CPR$ | 0.12 | 0.12 | 0.12 |
| Mean $\mathbb{E}[WAL]$ | 7.16 | 7.17 | 7.14 |
| RMSE (bps) | 27.22 | 23.29 | 32.44 |
| Mean $\alpha$ | 0.88 | 0.84 | 0.93 |
| Mean $\beta$ | 3.11 | 2.97 | 3.31 |
| Mean $\mathbb{E}[Loss]$ | 112.99 | 105.65 | 122.73 |
| SD $\mathbb{E}[Loss]$ | 44.84 | 42.94 | 45.86 |
| Min $\mathbb{E}[Loss]$ | 21.28 | 21.28 | 26.58 |
| Max $\mathbb{E}[Loss]$ | 228.87 | 228.87 | 220.23 |
| Mean $\mathbb{E}[g-fee]$ | 15.79 | 14.74 | 17.18 |
| SD $\mathbb{E}[g-fee]$ | 6.29 | 6.00 | 6.44 |
| Min $\mathbb{E}[g-fee]$ | 2.98 | 2.98 | 3.76 |
| Max $\mathbb{E}[g-fee]$ | 31.93 | 31.93 | 30.72 |

**Table 1.11:** Estimation Results: CPR of 12%

| | All | Group 1 | Group 2 |
|---|---|---|---|
| N | 121.00 | 69.00 | 52.00 |
| $CPR$ | 0.18 | 0.18 | 0.18 |
| Mean $\mathbb{E}[WAL]$ | 5.40 | 5.41 | 5.39 |
| RMSE (bps) | 27.82 | 23.99 | 32.90 |
| Mean $\alpha$ | 0.93 | 0.89 | 0.99 |
| Mean $\beta$ | 4.16 | 4.01 | 4.36 |
| Mean $\mathbb{E}[Loss]$ | 94.91 | 88.54 | 103.37 |
| SD $\mathbb{E}[Loss]$ | 40.00 | 38.43 | 40.82 |
| Min $\mathbb{E}[Loss]$ | 22.56 | 22.56 | 30.04 |
| Max $\mathbb{E}[Loss]$ | 201.02 | 201.02 | 191.98 |
| Mean $\mathbb{E}[g-fee]$ | 17.57 | 16.36 | 19.17 |
| SD $\mathbb{E}[g-fee]$ | 7.45 | 7.13 | 7.62 |
| Min $\mathbb{E}[g-fee]$ | 4.18 | 4.18 | 5.61 |
| Max $\mathbb{E}[g-fee]$ | 37.19 | 37.19 | 35.86 |

**Table 1.12:** Estimation Results: CPR of 18%

| | All | Group 1 | Group 2 |
|---|---|---|---|
| N | 121.00 | 69.00 | 52.00 |
| $CPR$ | 0.21 | 0.21 | 0.21 |
| Mean $\mathbb{E}[WAL]$ | 4.82 | 4.83 | 4.81 |
| RMSE (bps) | 27.86 | 24.15 | 32.79 |
| Mean $\alpha$ | 0.96 | 0.92 | 1.02 |
| Mean $\beta$ | 4.70 | 4.53 | 4.92 |
| Mean $\mathbb{E}[Loss]$ | 88.14 | 82.22 | 95.99 |
| SD $\mathbb{E}[Loss]$ | 37.98 | 36.63 | 38.65 |
| Min $\mathbb{E}[Loss]$ | 23.41 | 23.41 | 30.96 |
| Max $\mathbb{E}[Loss]$ | 189.64 | 189.64 | 181.07 |
| Mean $\mathbb{E}[g-fee]$ | 18.29 | 17.03 | 19.96 |
| SD $\mathbb{E}[g-fee]$ | 7.93 | 7.63 | 8.09 |
| Min $\mathbb{E}[g-fee]$ | 4.86 | 4.86 | 6.47 |
| Max $\mathbb{E}[g-fee]$ | 39.35 | 39.35 | 37.84 |

**Table 1.13:** Estimation Results: CPR of 21%

| | All | Group 1 | Group 2 |
|---|---|---|---|
| N | 121.00 | 69.00 | 52.00 |
| $CPR$ | 0.24 | 0.24 | 0.24 |
| Mean $\mathbb{E}[WAL]$ | 4.36 | 4.37 | 4.35 |
| RMSE (bps) | 27.83 | 24.24 | 32.59 |
| Mean $\alpha$ | 0.98 | 0.94 | 1.04 |
| Mean $\beta$ | 5.22 | 5.04 | 5.47 |
| Mean $\mathbb{E}[Loss]$ | 82.53 | 76.95 | 89.93 |
| SD $\mathbb{E}[Loss]$ | 36.22 | 35.00 | 36.82 |
| Min $\mathbb{E}[Loss]$ | 24.09 | 24.09 | 31.30 |
| Max $\mathbb{E}[Loss]$ | 180.35 | 180.35 | 172.03 |
| Mean $\mathbb{E}[g-fee]$ | 18.93 | 17.62 | 20.67 |
| SD $\mathbb{E}[g-fee]$ | 8.37 | 8.06 | 8.52 |
| Min $\mathbb{E}[g-fee]$ | 5.52 | 5.52 | 7.23 |
| Max $\mathbb{E}[g-fee]$ | 41.40 | 41.40 | 39.52 |

**Table 1.14:** Estimation Results: CPR of 24%

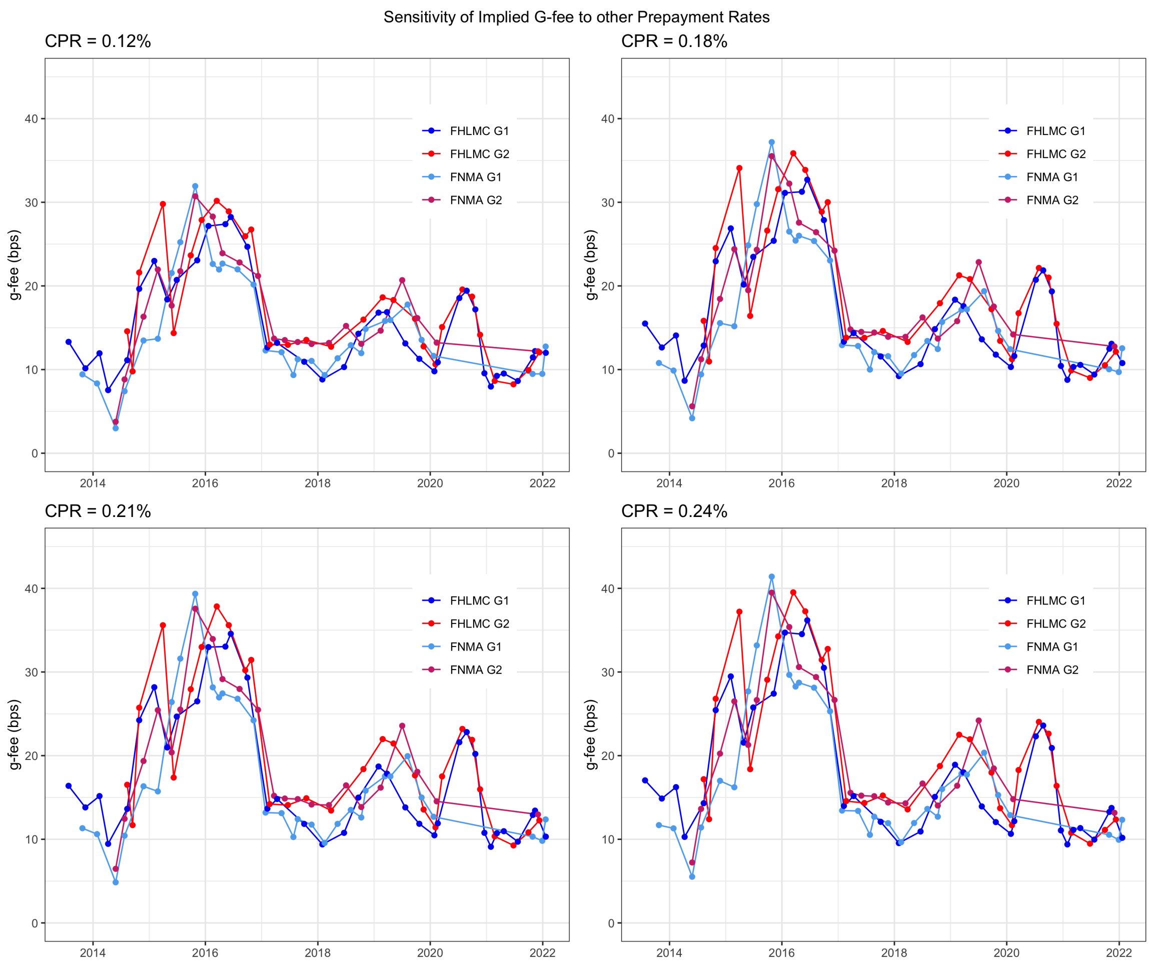



**Figure 1.18:** Robustness to Prepayment Speeds

### 1.A.8 Could GSE retained risk explain the g-fee spread?

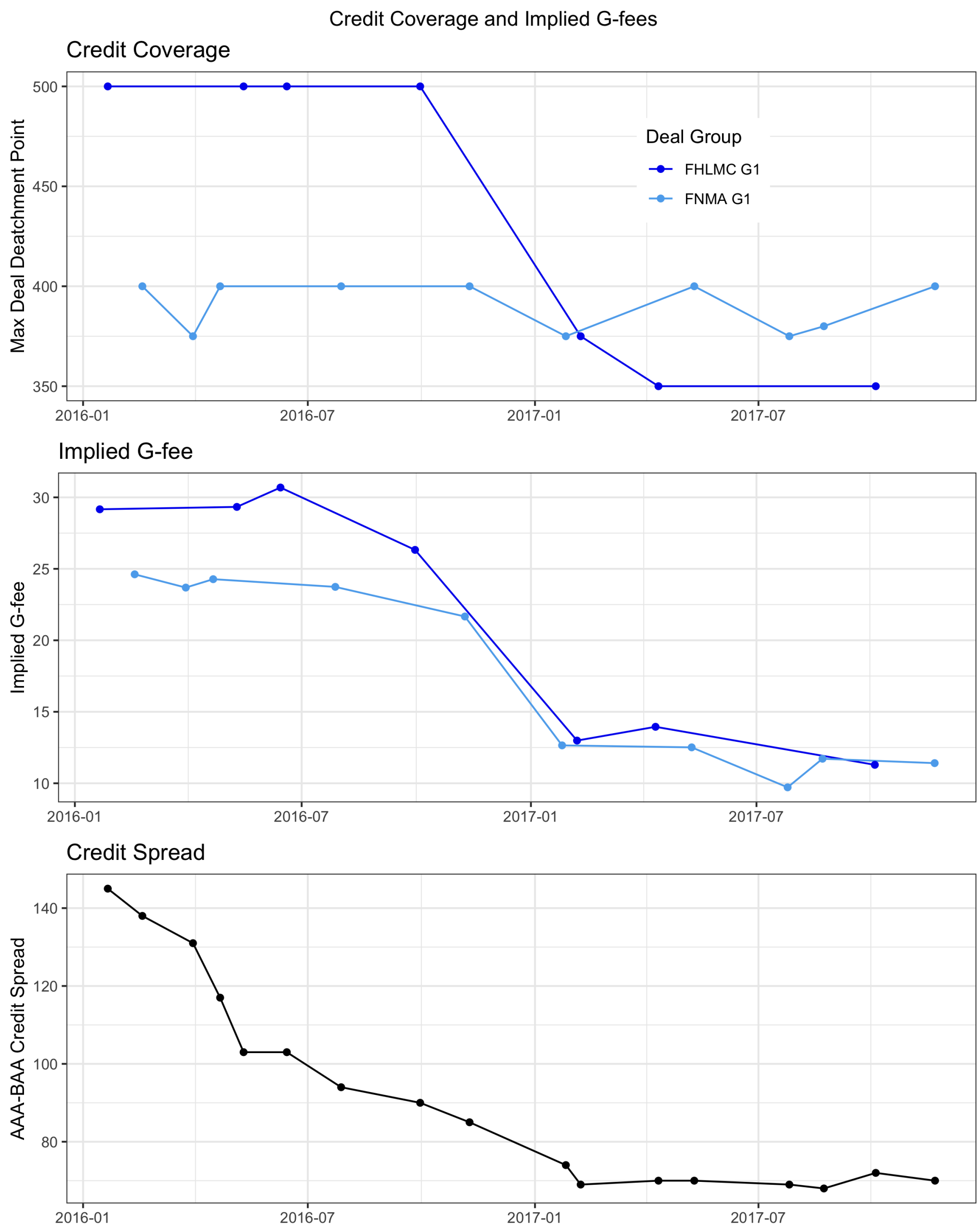



**Figure 1.19:** This map corresponds with the discussion on credit coverage the the implied measure of g-fees. In 2017, Freddie Mac brought the subordination levels of CRT deals inline with those of Fannie Mae. This provides a back-of-the-envelope estimate of the cost of retained risk to be around 5 basis points.

## References

Chernov, Mikhail, Brett R. Dunn, and Francis A. Longstaff. 2017. "Macroeconomic-Driven Prepayment Risk and the Valuation of Mortgage-Backed Securities." The Review of Financial Studies 31, no. 3 (December): 1132–1183.

Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin. 2001. "The Determinants of Credit Spread Changes." The Journal of Finance 56 (6): 2177–2207. Accessed May 10, 2022.

Cooperstein, Richard, and Michael Stegman. 2019. "A Missing Piece of the Administrative Reform Puzzle: How the GSEs Generate Cross-Subsidies." Joint Center for Housing Studies Harvard University (October).

Diener, Nicolas, Robert Jarrow, and Philip Protter. 2012. "Relating Top-Down with Bottom-Up approached in the Evaluation of ABS with Large Collateral Pools." International Journal of Theoretical and Applied Finance 15 (02): 1250011.

Diep, Peter, Andrea L. Eisfeldt, and Scott Richardson. 2021. "The Cross Section of MBS Returns." The Journal of Finance 76 (5): 2093–2151.

Echeverry, David. 2020. "Adverse Selection in Mortgage Markets: When Fannie Mae Sells Default Risk." Working Paper.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann. 2001. "Explaining the Rate Spread on Corporate Bonds." The Journal of Finance 56 (1): 247–277. Accessed May 10, 2022.

Fermanian, Jean-David. 2013. “A Top-Down Approach for Asset-Backed Securities: A Consistent Way of Managing Prepayment, Default and Interest Rate Risks.” The Journal of Real Estate Finance and Economics 46, no. 3 (April): 480–515.

Finkelstein, David, Andreas Strzodka, and James Vickrey. 2018. “Credit Risk Transfer and De Facto GSE Reform.” Economic Policy Review 24 (3): 557–595.

Fleckenstein, Matthias, and Francis A Longstaff. 2022. “The Market Risk Premium for Unsecured Consumer Credit Risk.” The Review of Financial Studies (January).

Gao, Chao, and John J. McConnell. 2018. “Investment Performance of Credit Risk Transfer Securities (CRTs): The Early Evidence.” The Journal of Fixed Income 28 (2): 6–15. ISSN: 1059-8596.

Gete, Pedro, Athena Tsouderou, and Susan M. Wachter. 2020. “Mortgage Stress without Government Guarantees. Lessons from Hurricanes and the Credit Risk Transfers.” Working Paper.

Gete, Pedro, and Franco Zecchetto. 2017. “Distributional Implications of Government Guarantees in Mortgage Markets.” The Review of Financial Studies 31, no. 3 (July): 1064–1097.

Giesecke, Kay, Lisa R. Goldberg, and Xiaowei Ding. 2011. “A Top-Down Approach to Multiname Credit.” Operations Research 59 (2): 283–300.

Golding, Edward, and Deborah Lucas. 2020. “Credit Risk Transfer and the Pricing of Mortgage Default Risk.” MIT Golub Center Working Paper.

Goodman, Laurie, Bob Ryan, Jim Parrott, and Mark Zandi. 2022. "How Fannie and Freddie Can Use Pricing to Expand Affordable Homeownership." Housing Finance Policy Center Commentary (April).

Goodman, Laurie, Ellen Siedman, Jim Parrott, and Jun Zhu. 2014. "Guarantee Fees - An Art, Not a Science." Housing Finance Policy Center Commentary.

Hayre, Lakhbir. 2001. "Salomon Smith Barney guide to mortgage-backed and asset-backed securities." Wiley and Sons 63 (2): 529–563.

Hurst, Erik, Benjamin J. Keys, Amit Seru, and Joseph Vavra. 2016. "Regional Redistribution through the US Mortgage Market." American Economic Review 106, no. 10 (October): 2982–3028.

Jeske, Karsten, Dirk Krueger, and Kurt Mitman. 2013. "Housing, mortgage bailout guarantees and the macro economy." Journal of Monetary Economics 60 (8): 917–935. ISSN: 0304-3932.

Lai, Rose Neng, and Robert Van Order. 2019. "Credit Risk Transfers and Risk-Retention: Implications for Markets and Public Policy." Working Paper.

Layton, Don. 2020. "Demystifying GSE Credit Risk Transfer: Part II - How, and How Well, Does It Work?" Joint Center for Housing Studies Harvard University.

Longstaff, Francis A, Sanjay Mithal, and Eric Neis. 2005. "Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market." The journal of finance 60 (5): 2213–2253.

Longstaff, Francis A., and Arvind Rajan. 2008. "An Empirical Analysis of the Pricing of Collateralized Debt Obligations." The Journal of Finance 63 (2): 529–563.

O'Neill, James. 2022. "A Top-down Model for Mortgage Portfolio Credit Risk." Working Paper.

Palmer, Kevin. 2017. "What Credit Risk Transfer Tells Us About Guarantee Fees." Freddie Mac, http://www.freddiemac.com/fmac-resources/perspectives/pdf/CRT-GFee-White-Paper-Final.pdf.

Richardson, Matthew P., Stijn Van Nieuwerburgh, and Lawrence J. White. 2017. "What to Do About the GSEs?" Annual Review of Financial Economics 9 (1): 21–41. https://doi.org/10.1146/annurev-financial-121415-033047.

Schwartz, Eduardo S., and Walter N. Torous. 1992. "Prepayment, Default, and the Valuation of Mortgage Pass-Through Securities." The Journal of Business 65 (2): 221–239. ISSN: 00219398, 15375374. http://www.jstor.org/stable/2353163.

Sirignano, Justin, and Kay Giesecke. 2019. "Risk Analysis for Large Pools of Loans." Management Science 65 (1): 107–121.

Vickery, James I, and Joshua Wright. 2013. "TBA trading and liquidity in the agency MBS market." Economic Policy Review 19 (1).

Wachter, Susan M. 2018. "Credit Risk Transfer, Informed Markets, and Securitization." Economic Policy Review 24 (3).

# CHAPTER 2

# A Top-down Model for Mortgage Portfolio Credit Risk

## 2.1 Introduction

In this paper, I offer an empirical analysis that extends previous work in O'Neill (2022) on the informational content of a new asset class, GSE Credit Risk Transfer (CRT) bonds. CRT bonds back the credit risk of mortgage loans purchased by the two Government Sponsored Enterprises (GSEs), Fannie Mae and Freddie Mac. When mortgage loans backed by CRT bonds suffer losses, principal is written down and the GSEs are reimbursed for losses by CRT investors. The intuition of the paper above was that the present value of losses incorporated in tranche market prices is revealing about the portfolio loss probability density function. This probability density function is the implied risk-neutral loss density function. This paper specifies a richer default process in order to make deeper inferences about how the market views credit risk in diversified mortgage pools. In the spirit of a reduced-form credit model, defaults are governed by an exogenous Poisson process. The particular specification I will use in this model uses two factors, which are the intensities associated with two different default risk factors.

Specifically, I estimate a reduced-form two factor no-arbitrage model of portfolio risk for conforming mortgages, to learn about how the market views credit risk in the underlying mortgage loans. In this model, mortgage amortization and prepayments are modeled in a top-down framework. This means that rather than model cash flows at the loan level, I take a statistical approach and model cash flows at the pool level. This simplification, which has been used in mortgage-backed securities modeling as well as in the valuation of portfolios of other credit sensitive instruments, is justified when there is substantial homogeneity in the underlying constituent loans (Diener, Jarrow, and Protter (2012)). This is precisely the case in GSE mortgage pools and CRT collateral; CRT bonds are backed by substantially similar single-family 30 year loans originated around the same time with similar interest rates.

The principal contribution of this paper is to offer one of the simplest possible reduced-form credit models for estimating parameters of interest related to the market pricing of mortgage credit risk, which was not possible until the advent of the GSE credit risk transfer programs in 2013, when conforming mortgage credit risk began to be traded in secondary markets. In estimating the model, I provide estimates of market implied default probabilities and the market-view of the probability of housing market crises. It is the first paper to my knowledge to quantify how the market views the probability of an adverse event hitting the conforming mortgage market. Due to the systemic importance of the housing market, this model could be used as a real-time risk assessment for the level of risk in the mortgage market, similar to the way that other derivatives markets use model based measures of implied volatility, tail risk or correlations. In addition to this new model, I simulate mortgage pay downs and defaults and demonstrate what the results mean for housing market risk and the design of CRT bond programs.

The model's two factors, $\lambda_1$ and $\lambda_2$, are estimated for 32 Freddie Mac CRT issuances from 2018 through 2022. The factors are motivated by historical default behavior in the Fannie Mae Single-Family Data set.[1] Figure 2.1 below provides this

1. Since GSE loans are substantially similar between Freddie Mae and Fannie Mae, I do not make a distinction in the model or in the data. This reasonable assumption is supported by the fact that FHLMC and FNMA are now deliverable in the same TBA mortgage-backed security, the UMBS

motivation by comparing overall loss rates on loans originated between 1999-2001, cohorts considered to have suffered relatively benign levels of losses, and 2005-2007, loans that suffered severe losses during the Global Financial Crisis. Although this is not an exhaustive plot of losses on mortgage originations, it is suggestive that loan losses are bimodal in the sense that low, relatively predictable levels of losses are incurred in good times and that losses can be several orders of magnitude higher when a crisis strikes the housing market; losses range from around 25 basis points to over 300 basis points during the worst years of the global financial crisis (GFC). The two factor model formalizes the intuition that the market prices both types of risk in the CRT market.[2]

(FHFA Single-Security Initiative).

2. I provide evidence that a one-factor model performs substantially worse, with pricing errors 3-15 times larger on average in section 2.5.4.

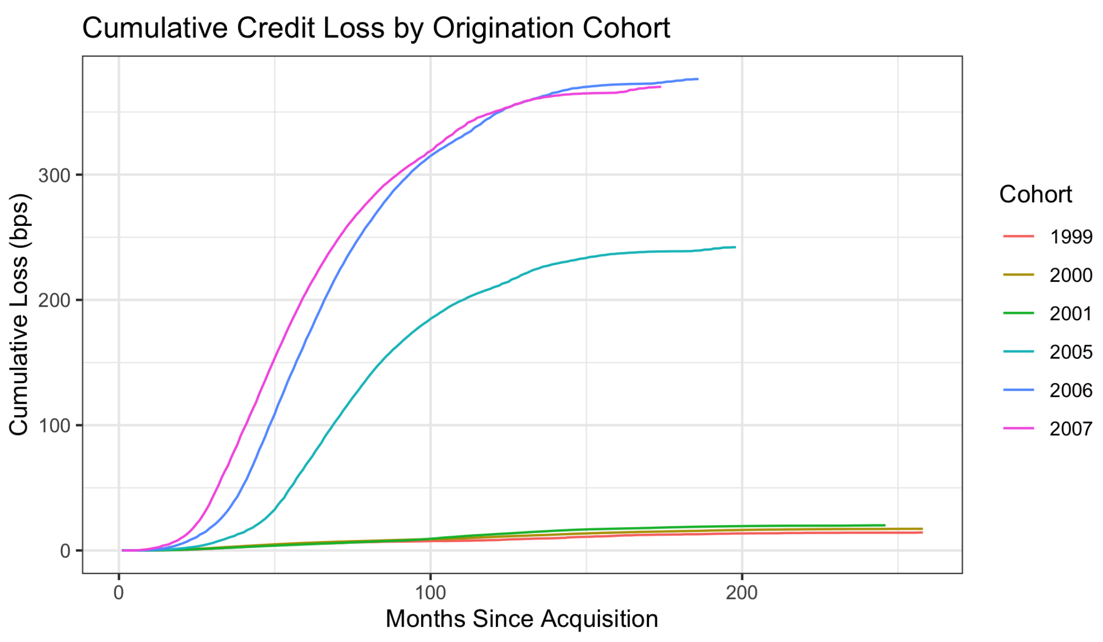



**Figure 2.1:** Cumulative losses for selected historical FNMA mortgage cohorts. These cohorts are filtered to include only loans that have LTVs between 60 and 97 and credit scores between 640 and 780, to make them most comparable to the CRT collateral.

Further motivation is provided in the selection of parameters for the model, discussed in section 2.4, where I show that default rates, time to liquidation, and eventual loss severity differ between these two "types" of default behavior. Loans originated leading up to and during the crisis period experienced longer default timelines and higher severities when the housing market suffered. Appendix sections 2.A.5, 2.A.6, and 2.A.7 provide further historical data on the different nature of default rates, recovery rates, and time to liquidation in both normal and crisis periods. Including these features in the model is important because losses are not immediately realized on CRT bonds when defaults occur; investors incur losses when the GSE writes-down the loans after the default process has concluded.[3]

The two factors map naturally into this intuitive understanding of credit risk in the mortgage market. The first-factor represents a routine level of defaults; a jump size of 0.0001 on a notional pool scaled to \$1 corresponds to 10 loans in a typical CRT backed mortgage pool of $100,000$ loans entering default in a given month. The estimated Poisson intensity of this risk factor is 1.897 for the sample of CRT bonds studied herein, suggesting that in a given month the market expects about 19 loans out of 100,000 to enter default. Put differently, the risk-neutral expectation is that around 2 basis points of loan principal are defaulted upon in a given month. This means that the market expects routine default levels roughly twice as high as in the historical data. This is an interesting finding because default levels during calm housing markets are generally considered easily predictable given loan level characteristics. CRT prices suggest investors may require a risk premium on both components.

The second risk scenario represents a housing crisis situation in which 3.2% of loan principal enters default. This number corresponds closely to yearly default levels during the peak of the financial crisis in 2007 (See section 2.4 and appendix

3. See Finkelstein, Strzodka, and Vickrey (2018) and O'Neill (2022) for further discussion on the structure and economics of CRT bond deals, including example paydowns.

section 2.A.5). This factor could be interpreted as capturing the higher correlation, or clustering, between defaults in a crisis scenario. The mean model estimated market implied risk-neutral probability of this event occurring in a given month is 0.0035, or once every 286 months (24 years). This probability may come as a surprise given the increased under writing quality in GSE loans since the crisis and the implications of such an event. Estimating such an event's probability under the physical measure $\mathbb{P}$ is beyond the scope of this paper, but it provides suggestive evidence that the market does not view such an event as particularly unlikely and that private capital places a substantial risk premium on the pricing of catastrophic mortgage credit risk.

The top-down framework employed in this paper offers a substantial simplification over having to specify a full econometric model for predicting mortgage defaults; it requires vastly fewer parameters to calibrate. It also differs in another crucial dimension; because resulting probabilities in the model are estimated under the risk-neutral measure, it can speak to the market pricing of credit risk. Simulation-based econometric models of mortgage default can provide accurate predictive modeling, but tell us nothing about risk-premiums in the the mortgage credit market. It would instead require the use of an option-adjusted spread (OAS) to match tranche market prices. Since each tranche would require its own OAS, it is not clear that the no-arbitrage restriction that requires each tranche to derive its value from the same underlying risk-neutral loss distribution is maintained (Gauthier (2003)).

Despite its simplicity compared to full-scale econometric models calibrated under the $\mathbb{P}$ measure, the model is general enough to allow default behavior to differ upon the realization of a default: in this top-down framework, a default realization occurs upon a Poisson event and results in a fixed portion of loans entering default. After a period of time, a loss is realized on a portion of the loans based on a loss severity. Of course in reality, a portion of loans may default but subsequently become current with payments and therefore not realize a loss. In this top-down model, such behavior would be incorporated into loss severities. For example, if 1% of loans enter default in a given period and the loss given default is 30%, one could interpret the recovery of 70% of the loan balances as including the recovery of the present value of loans who continue to make payments. The top-down model need not distinguish between these two for the sake of valuation. This vastly simplifies the specification of a bottom-up mortgage default model which could require hundreds of parameters if not more.

Furthermore, the model could be expanded greatly to allow for randomness in jump sizes, loss severities, and the time between loans entering default and the subsequent write down (loan disposition). Although I focus on the case in which two default intensities are the driving state variables, there is not reason why the model could not be extended to the case with $N$ default risk factors, or stochastic default intensities. I argue in this paper using historical data on defaults that the two principal types of default scenarios provide justification for why the two factor model performs well.

This richness is bounded by computational complexity as well as ensuring the identification of the parameters of interest. I restrict my sample to actual-loss CRT bond issuances that include four or more tranches, and estimate only two state variables per issuance in order to ensure that the model is restricted and tranche prices are not being over-fit. Despite this conservative restriction, I find that the model fits tranche prices well over the 32 bond issuances in the sample, with an an average tranche level root mean squared error of 97 basis points. Future research may determine which key variables should be estimated using market prices and which should be taken from historical data.[4] For simplicity, I do not estimate the full set of model parameters, but take parameters as given as well motivated by historical data and

4. The caveat, of course, is that there is no guarantee that level or dynamics of a particular parameter under $\mathbb{Q}$ have to look like dynamics under $\mathbb{P}$. That being said, we may be able to make reasonable assumptions about where functions are likely to differ under the two measures (i.e., carry a risk premium).

solve for the associated levels of the state variables.[5]

Additionally, implied guaranty-fees (g-fees) can be estimated from the results in this paper. A mapping of the into a g-fee shows that the market considers fair compensation for bearing this credit risk to be on the order of 10-20 basis points a year, which compares to the roughly 30 basis points charged by the GSEs for insuring it, consistent with the results of the model in O'Neill (2022) which estimates loss distributions at the maturity of the mortgage pool only. Again, since the guaranty-fee explicitly incorporates a risk premium component, models that are not based on actual tranche market prices cannot speak to the full market implied g-fee, but rather only the expected loss component. I also provide a decomposition of the implied guaranty-fee and show that the market's required compensation for the first risk factor has increased of late, therefore increasing its share of the implied g-fee.

Lastly, I use the results mentioned above to make some comments on the credit protection that CRT bonds offer the GSEs. I find that in most circumstances, CRT bonds do provide ample credit protection despite their complicated paydown structure. In 10% of paydown scenarios, the GSE takes losses 10-20% of the time, which is a function of both the levels of the risk factors and the detachment point of the CRTs subordinate balance. This paper, being the first to systematically price CRT tranches and explore the informational content of these bonds, provides future research directions and modeling guidance. An increase in CRT bond issuance, as well as the expansion of different types of CRT bond capital structures offered, could aid researchers in further disentangling how the market views mortgage credit risk. This paper serves as the first step in that direction.

The rest of the paper proceeds as follows: I briefly review the related literature

5. In the paper, the phrases risk factors, state variables, levels of the default process, and default intensities all refer to the same thing. The choice of terminology is usually determined by the context in which the model is being discussed.

below in section 2.1.1 and the data in section 2.2. In section 2.3 I introduce the two factor valuation model. In the remainder of the paper I focus on the empirical applications of the model. I provide further motivation for the modeling choices in section 2.4, by focusing on the historical dynamics of mortgage defaults in both normal times and crisis episodes. The estimation strategy and results are discussed in section 2.5, and the implications for the results in section 2.6. Lastly, in section 2.7 I conclude, and offer direction for future research on the subject of GSE credit risk transfers.

### 2.1.1 Literature

This paper, along with O'Neill (2022), is the first to my knowledge to apply a reduced-form credit model framework to the study of CRT bond tranches as well as the first attempt to use actual CRT market prices to derive a market view of mortgage credit risk. In these reduced-form credit models, default dynamics are governed exogenously, usually in the form of an intensity process. An early example can be found in Pye (1974), which derives market-implied default probabilities from bond yields. Duffie and Singleton (1999), Jarrow and Turnbull (1995), Litterman and Iben (1991), and Duffie and Garleanu (2001) are influential examples of these types of models in the context of risky corporate debt. The main advantage of these models are calibration to market prices and their use in hedging and risk management. They are also especially useful when features related to underlying credit riskiness are difficult to measure or unobservable, and useful when defaults are correlated, as in the case with mortgage pools. Bluhm and Wagner (2011) provide a summary of applications for portfolio credit risk.

I combine these notions under the umbrella of top-down credit modeling, in which cash flows from individual portfolio constituents are aggregated to describe paydown dynamics in a parsimonious way. Giesecke, Goldberg, and Ding (2011) and Longstaff and Rajan (2008) apply top-down models to portfolios of corporate credit. Fermanian (2013) is one of only a few examples of applying this method to risky mortgage debt, in the context of CMOs. Giesecke, Goldberg, and Ding (2011), Diener, Jarrow, and Protter (2012), and later Sirignano and Giesecke (2019) provide evidence on the conditions under which top-down models perform well as approximations for large pools of risky constituent assets. These papers demonstrate that these models can provide rich economic implications while avoiding much of the computational complexity associated with full-scale econometric models. Most importantly, these models are fully calibrated under the risk-neutral measure and thus can directly spread to risk premiums on multiple credit dimensions if the $\mathbb{P}$ measure counterpart can be directly observed. Chernov, Dunn, and Longstaff (2017) provides an example of this in the context of mortgage prepayment risk.

Most importantly, this paper contributes to the debate on GSE reform by making progress on the extent to which mortgage credit information can be extracted from opaque CRT securitizations. Much of the recent literature related to CRT bonds is focused on their qualitative features and their relationship to GSE mandates on diversifying risk. Wachter (2018) and Finkelstein, Strzodka, and Vickrey (2018) examine the structure of the bonds and their relationship to the GSEs market model.

Few papers have yet to examine the pricing of the bonds: Gao and McConnell (2018) examines the realized returns on early CRT issuances, Belbase (2014) examines the impact of several stress scenarios on tranche performance, and Golding and Lucas (2020) simulates the paydown on an example tranche. I contribute to this literature by building and estimating a portfolio credit model matched to actual tranche prices using a comprehensive data set on recent CRT issuances by Freddie Mac.

## 2.2 Bond Data

To fit the pricing model, I use the data set of O'Neill (2022) which contains comprehensive deal and tranche level information for GSE issuances from both Fannie Mae and Freddie Mac, as well as average tranche market prices after issuance from TRACE. I choose to estimate the pricing model in this paper on only a subset of the original dataset which contains 32 bond deals from Freddie Mac and 129 tranche prices.[6] These bond issuances are referred to as "STACR" bonds (Structured Agency Credit Risk), and are issued in two groups related to the loan-to-value ratios of the underlying loans. Group 1 deals feature underlying loans with LTVS of 60-80%, and Group 2 deals feature collateral LTVs of greater than 80% to the conforming limit of 97%.

There are several main reasons why only a subset of the full CRT issuance data set is used. First and foremost, CRT issuances have become standardized as of late to include 4 tranches that provide more consistent credit coverage than some of the earlier issuances. Credit coverage begins around 15 to 25 basis points of losses, which helps identify the state variable $\lambda_1$. Earlier CRT issuances did not issue junior bonds, therefore this state variable would have no impact on the root-mean-squared error (RMSE) of trying to fit tranche prices and is thus fundamentally unidentifiable, meaning that earlier deals only spoke to "disaster" risk in mortgage markets; I find in this paper that both types of risk appear important to the pricing of CRT bonds. Second, because there are 4 issuances, the model is sufficiently overidentified, requiring the model to fit 4 tranche prices with 2 state variables.

Third, all of the bonds featured in the data set here are actual loss bonds. This means that there is no discrepancy between the actual loss realized on the under-

6. The latest issuance, the STACR 2022 deal, has 5 tranches, which is a promising sign moving forward that CRT deals will continue to cover more of the capital structure.

lying mortgage loan and the write-down on the CRT bond. This removes an extra assumption from being made, in particular that the loss tables used on early "fixed-severity" CRT bonds accurately reflects the eventual losses that will be incurred on the bonds. A further institutional feature that is largely consistent throughout this sub-sample is the REMIC designation, which was a significant feature implemented in 2019 that improved the tax treatment of STACR bonds and therefore increased the pool of potential investors in CRT bonds and therefore their liquidity in the secondary market.[7]

There is still heterogeneity in the bonds that are in this sub-sample. For example, Freddie Mac STACR bonds in this sample have a maturity that varies between 12.5, 20 and 30 years. The total level of subordination varies as well, with STACR subordination levels coming down during the COVID-19 pandemic before reverting back to normal pre-crisis levels of about 4%.[8] Appendix table 2.3 shows the bond issuances included in the sample, along with the number of tranches offered. Also included is the weighted average coupon of the underlying mortgages, along with the weighted average maturities, credit scores, and loan to value ratios.

Table 2.4 shows the market prices for each tranche, listed by seniority from left to right in the table, with M1 referring to the most senior subordinated tranche and B2 referring to the most junior tranche. For conciseness, I do not repeat in depth discussion of summary statistics that can be found in O'Neill (2022). For reference, the tranche attachment and detachment points can be found in the appendix plot 2.A.3, and the tranche floater spreads can be found in appendix table 2.A.4. Next, in section 2.3, I formally describe the two factor model and the functional form that will be specified for defaults.

7. See the Freddie MAC CRT Handbook for further information on these features.

8. See Netter (2020) for a discussion on the CRT market in relation to bond market disruptions during 2020.

## 2.3 Valuation Model

The inherent complexity of Credit Risk Transfer securities may explain the relative lack of interest in their pricing and performance in the academic literature. Nevertheless, they are the first chance that researchers have had to observe market pricing for conforming mortgage credit risk. Furthermore, they have improved liquidity and transparency vis-á-vis previous generations of private label MBS and CMOs. Debate on housing market reform in the United States must be centered on evidence pertaining to the pricing of credit risk by private market participants; the CRT bonds provide a compelling place to start this debate.

The valuation of CRT bonds is a high-dimensional problem, and certain assumptions will have to be made to render their valuation, and the extraction of implied default rates, tractable. I now formally describe the valuation framework, starting with a quick refresher on the structure of CRT tranches and the intuition for the top-down modeling convention and then next, the two factor model of defaults.

### 2.3.1 CRT Bond Overview

CRT tranches represent a form of synthetic credit derivative on the performance of loans in the underlying reference pool. In order to maintain the benefits of the highly liquid TBA market for mortgage-backed securities for both the investors and the GSEs, cash flows on CRT bonds are independent of the actual cash flows received from the paydown of the mortgages themselves; instead, their cash flows simply reference the performance of the loans. The GSEs retain the senior most portion of the capital structure of the CRT bonds, often around 96%. The junior portion is sliced into tranches with differing credit risk attributes and cash flow seniority.

Cash flows, in the form of scheduled principal and principal prepayments are allocated pro-rata between the senior share and the junior subordinate portion. Within the subordinate portion, principal payments are paid sequentially with more senior bonds being paid first. Defaults are allocated in the opposite direction, with the most junior bonds having their principal written down first upon realized losses on the principal. While each CRT tranche still has principal outstanding, investors receive a coupon payment based on the prevailing LIBOR/SOFR rate and a floater spread, which can range from around 70 basis points to over 1200 basis points depending on seniority (see appendix figure 2.9). For a more in-depth treatment on the structure of CRT bonds, their role in the GSE market model, and examples of their cash waterfalls, see Finkelstein, Strzodka, and Vickrey (2018), O'Neill (2022), Belbase (2014), and others.[9]

### 2.3.2 Top-down Valuation Framework

As mentioned above, the model combines the reduced-form approach with a top-down specification for principal paydowns and defaults, in which default dynamics are modeled on the whole pool of diversified loans. This is similar to the way that mortgage pool prepayments are described using the Conditional Prepayment Rate (CPR), which represents a sufficient statistic for prepayment behavior in the aggregate. By extending this notion to defaults, the goal is to remain as agnostic as possible about the potential drivers of default, but rather describing the process of state variables that allow us to best match market prices of credit-linked bonds. Examples of top-down models being applied to mortgage bonds include Fermanian (2013) and Chernov, Dunn, and Longstaff (2017).

The notion of top-down credit modeling is straightforward in the context of CRT bonds. In a portfolio of diversified loans, there is a Poisson intensity variable $\lambda$ that governs the arrival of default on a particular portion of the loan principal. In order

9. Prospectuses for CRT deals can be found on the GSE websites: Fannie Mae, Freddie Mac

to have as few free parameters as possible, as there are only four tranche prices per CRT bond issuance, I assume that intensities are non-stochastic and that there is a constant intensity that best matches the cross-section of tranche prices for each bond issuance. The model does not collapse into the simple sum of two Poisson processes because each risk-factor is associated with its own liquidation timeline and eventual severity, which govern the actual losses realized on CRT tranches and therefore their prices.

In order to follow the cash flows linked to CRT bonds and the valuation procedure, I proceed in the following steps: First, section 2.3.2.1 describes the cash flows from the underlying mortgage pool, which consist of scheduled principal, unscheduled principal (prepayments), and defaults. Section 2.3.2.2 describes the formal modeling the defaults in the two factor specification, which leads to the tranche valuation expression in section 2.3.2.3. To complete the model, I describe prepayment rates and choice of the discount rate function in sections 2.3.2.4 and 2.3.2.5.

#### 2.3.2.1 Modeling Mortgage Cash Flows

Consider a continuous time setting where a diversified pool of homogeneous mortgages each have maturity $T$ and pay a continuously compounded interest rate $r$. The starting total principal balance is normalized to equal \$1. Standard formulas apply for the monthly mortgage payment along with its principal and interest components. To distinguish between the notional balance of mortgage principal remaining and the actual balance of the mortgage principal remaining, I define the survival factor $Q_t$. $Q_t$ represents the total fraction of pool principal that has yet to be defaulted upon or prepaid (since the loans are homogeneous, you could also interpret this as the fraction of loans in the hypothetical that the pool was made up of infinitely many loans).[10]

10. See Hayre (2001) for more formulas and examples of this type of mortgage math. The Bond Market Association also publishes standard formulas here.

Of course, we care intimately about the composition of $Q_t$ between loans that have prepaid and loans that have defaulted for the purposes of properly allocating cash flows for CRT bonds: prepayments simply return capital to investors where as defaults result in a write-down of investor principal and a payment to the GSEs for reimbursement. To do so, I define further pool factors $P_t$ and $D_t$ which represent the cumulative contributions to $Q_t$ that come from prepayments and defaults, respectively. $P_t$ can be expressed as $P_t = exp(\int_0^t -p_t dt)$, where $p_t$ is the instantaneous prepayment rate. Intuitively, this is like the price of a zero-coupon bond with a continuously compounded interest rate. $P_t$ can also be interpreted as the total percentage of pool notional that has yet to be prepaid in the absence of defaults; in the presence of defaults, these amounts will differ because the pool will be amortized by defaults at the same time that loans are prepaying.

$D_t$ is defined analogously to $P_t$, where instead $d_t$ is the instantaneous default rate. In the model, defaults and prepayments are assumed to be independent. This leads to the following formula for the actual pool principal balance at time $t$, $B_t^*$:

$$B_t^* = B_t \times Q_t = B_t \times P_t \times D_t \tag{2.1}$$

$B_t$ represents the pool balance the absence of prepayments or defaults and simply represents summarizes scheduled amortization; scaling by $Q_t$ gives the value in the presence of these two additional effects. To calculate the actual dollar amount of prepayments or defaults, we would take the prepayment or default rate, $dP_t$ or $dD_t$, and multiply by $B_t^*$.

This algebra ensures that the paydown of the mortgage pool is calculated properly while also making sure that the common terminology related to prepayment and default rates can maintain their usual interpretations. This will become more apparent as I discuss how the default rate is calculated in the credit model and the estimation results. In the following section, I describe the exact form that $D_t$ will take in the model and how it will map to a notion of mortgage credit risk.

#### 2.3.2.2 Modeling Defaults

As described above, $d_t$ represents the default rate, or intensity, in a given instant. A positive $d_t$, however we should choose to specify it, will eliminate principal from the underlying loan pool proportional to the default rate. It is helpful to consider an example: consider a time period $t$ by which the scheduled paydown of the loans has resulted in $B_t = 0.8$. This is the principal remaining on a loan with rate $r$ and maturity $T$; $P_t = 0.9$, and no defaults have occurred. If in the next instant, $d_t = 0.03$, ($B_t \times P_t \times D_t = 0.8 \times 0.9 \times 0.97 \times 0.8 = 0.698$), 69.8%, of the total original pool principal remains as performing. Additionally, it can be seen that 2.2% of the original pool balance has entered default. Note that this has the implication that for a given default rate, the paydown of the pool from scheduled amortization as well as principal repayments will effect the amount of total principal, in dollar terms and in fraction of the pool, that is actually defaulted upon.

I have now defined $d_t$, from which $D_t$ follows, the objects that tracks the level of default in the pool. The contribution of this paper is to provide a simple two factor model of $d_t$ and apply it to the valuation of CRT tranches in order to decipher how the market views credit risk; the CRT valuation problem effectively reduces to estimating the distribution of paths that $d_t$ takes. I specify a two factor model of portfolio credit risk in which the instantaneous default rate $d_t$ is driven by two Poisson intensities, $\lambda_1$ and $\lambda_2$. Define jump sizes $\gamma_1$ and $\gamma_2$, which represent the shock to the default intensity upon the realization of a default event, Poisson variable $N_{it}$:

$$d_t = \gamma_1 dN_{1t} + \gamma_2 dN_{2t} \tag{2.2}$$

Plugging into the definition of $D_t$ above:[11]

$$D_t = exp(-\int_0^t (\gamma_1 dN_{1t} + \gamma_2 dN_{2t})) \tag{2.3}$$

Integrating we get the result that $D_t$ is equal to:

$$D_t = exp(-\gamma_1 N_{1t} - \gamma_2 N_{2t}) \tag{2.4}$$

Several intuitive conditions are specified by this specification. The default survival factor, $D_t$, is bounded between 0 and 1, and is by definition equal to 1 at time 0. Its interpretation in the mortgage cash flow equations is preserved as it falls between 0 and 1 to represent that at no time more than 100% of the loans in the pool can enter default. Said alternatively, the total balance of loans which has been defaulted upon is a non-decreasing function of $t$.[12]

The most pressing modeling question at this stage is how precisely the defaults specified by the two factor model translate into loss write-downs on the pool of mortgages. Once default is triggered on a portion of pool principal, the amount is now no longer included in $B_t^*$, meaning that scheduled principal payments are no longer made on that amount nor can that portion be defaulted upon again or prepaid.

I must now specify a loss function that maps from defaults to actual loss write-downs on the CRT tranches. To do so, I introduce 4 more parameters into the valuation model. These parameters $\Theta$ differ from the state variables $\lambda_1$ and $\lambda_2$ in

11. Another option would be a linear specification, in which $D$ could theoretically fall below 0 but would be unlikely under any reasonable specification of the parameters.

12. For simplicity, I assume that the intensities are not stochastic. The problem cannot be reduced as the sum of two Poisson processes because the realization of each particular type of default risk has different implications for the losses realized on the mortgage pool, as is described below.The model maintains enough generality so that one could estimate stochastic intensities as a way to insert further default correlation.

that they are invariant between different bond issuances. This is to restrict the model, which would have more parameters than tranche prices if I allowed $\Theta$ to take on different values for each bond series issuance date. $\Theta \in [lgd_1, lgd_2 del_1, del_2]$, where $lgd_i$ and $del_i$ represent the severity of loss in the case of each type of risk realization and the time between default and loss write down (this is called many different things in the literature on mortgage default: time to liquidation, for example), respectively. Since $N_1$ and $N_2$ are independent Poisson random variables, we can separate the default risk from each state into separate components in order to identify the losses incurred from each type of risk. The balance of loans that enter default, $def_t$ is equal to:

$$def_t = B_t^* \times d_t = B_t^* \times [\gamma_1 dN_{1t} + \gamma_2 dN_{2t}] \tag{2.5}$$

There are a few important things to remember about $def_t$. First, it is an actual dollar amount rather than a proportion or an intensity such as $d_t$. Second, it be can decomposed into "type 1" defaults, and "type 2" defaults, or defaults caused by the relevant Poisson intensity $i \in [1, 2]$:

$$def_{i,t} = B_t^* \times \gamma_i dN_{it} \tag{2.6}$$

By the model definition, losses occur at time $t + del_i$ at a severity of $lgd_i$. Renormalizing to the loss period rather than the default period:

$$dl_{i,t} = B_t^* \times lgd_i \times \gamma_i dN_{i,t-del_i} \tag{2.7}$$

The total loss given in instant $t$ is the sum of the two constituent losses, $l_t = l_{1,t} + l_{2,t}$, and the cumulative losses are $L_t = \int_0^t dl_t$, which is bounded at time $t$ by virtue of being a function of $B_t^*$ (i.e., the loan losses cannot exceed the remaining principal balance).

The endless notation is necessary to preserve the distinction both between the default rate and the total notional defaulted, and the distinction between the notion of entering default and actually realizing a loss. To reiterate, the Poisson state variables $\lambda_1$ and $\lambda_2$ drive variation in the default rate - but the actual paydown of the pool, information that is summarized in $B_t^*$ and is an outcome of both defaults, prepayments, and scheduled amortization, is necessary to calculate actual dollar loss amounts. Dollar loss amounts, of course, are directly mappable to the write down on CRT tranches.

Time to liquidity and loss severity are important features that are almost always included in a model of mortgage default. These two features of the model help better match the reality of mortgage defaults, in which a total loss of principal rarely ever occurs (due to the lending being secured by the underlying property), and the bankruptcy proceeding taking some amount of time to play out. Section 2.4 goes into more detail about the historical dynamics of mortgage defaults and the role these two parameters play in reality.

I note that this is among the simplest possible two factor specifications that still contain these features. Researchers in the future may be interested in adding stochastic Poisson intensities, stochastic recovery rates/loss severities, and correlations between these variables.[13]

13. Longstaff and Rajan (2008), in their model of CDOs, note that a intensity that trends downward may reflect low quality firms exiting the portfolio and general credit risk improving over time. A mortgage model may feature the opposite dynamic - low quality loans do not prepay and thus sit around until they default later in the life of the pool.

#### 2.3.2.3 Tranche Cash Flows and Valuation

The repayment of principal on the bonds is allocated pro-rata between the senior and subordinate portions; within the subordinate portion, principal cash flows are allocated by seniority. Thus, as write-downs occur on the junior portion of the bonds, credit enhancement of the senior portion is reduced and more principal flows to the retained portion. Therefore, these bonds fit at least generally into the class of sequential pay CMOs that became popular in the 1990's.

A continuous time version of the formulas described in equations 4 and 5 of O'Neill (2022) can be used to determine the dynamic principal balances; since the model will be solved by discretizing the cash flow formulas and using Monte Carlo simulation, I opt to simply use these formulas directly.

Formulas for the cumulative principal and loss claims are found in equations 8 and 9 of O'Neill (2022). As is common on these types of bonds, these cumulative claims look like call spreads on the underlying loan balance and accumulated loan losses.[14] One interesting adjustment is that the upper strike of the principal claim call spread is moving downwards as defaults cause write-downs on the tranche; a tranche can never claim more principal than its remaining balance.

To tie together the final valuation expression, we define a CRT bond tranche with attachment and detachment points $K_j$ and $K_{j+1}$, and a principal balance of $B_{K_j,t}$. The principal balance of $B_{K_j,t}$ is the result of scheduled amortization, principal prepayments, and loss-write downs occurring as the result of default events that arrive at the Poisson state variable intensities.

The value of the tranche is thus given by the sum of present value of the coupon

14. As common in many other structured loan products, CRT bond cash flows may vary slightly according to the performance of the underlying loans. Appendix section 2.A.12 outlines some of these alternate features and why I do not believe they have a material effect on the results in this paper.

payments paid at floater spread $s_{K_j}$ above LIBOR/SOFR rate $q_t$ , the present value of the returned principal payments ($PC_t$ representing the cumulative principal claim), and the remaining balance to be repaid at maturity $T$. With $r_s$ is the short rate process, the valuation equation for a tranche with balance given by $B_{K_j,t}$:

$$PV(K_j, T) = E_t^Q[\int_0^T e^{-\int_0^t r_s ds}[(q_t + s_{K_j})B_{K_j,t} + PC_t]dt + \underbrace{e^{-r_T} B_{K_j,T}]}_{\text{Principal at Maturity}} \tag{2.8}$$

Again, there is assumed to be no correlation between prepayments, defaults, and the interest rate.

#### 2.3.2.4 Prepayments

Part of completing the model includes specifying a process for mortgage prepayments. Since the main focus of this paper is on mortgage defaults, not prepayments, I opt for one of the simplest specifications of prepayments possible while still respecting the fact that prepayments are random. On a particular valuation run, a prepayment rate is drawn at random from a uniform distribution with bounds of 12 and 28%, expressed in constant prepayment terms (CPR).

$$\forall t, p_t \sim Unif(\underline{p}, \overline{p}) \tag{2.9}$$

In a similar vein, the level of notional at which loans stop prepaying is drawn at random from a uniform distribution with bounds of 5 and 10%. This feature is included in order to mimic the burnout seen in the prepayment of mortgage pools. This means that eventually, loans stop prepaying no matter what. This feature is important because, without it, high prepayment rates mean that the entire pool notional could pay down quickly which is empirically implausible, and would make the CRT tranches appear unrealistically "safe."

#### 2.3.2.5 Coupon Rates and Discount Rate Considerations

To complete the expression, I need an approach for estimating the discount function, $D(t) = e^{-\int_0^t r_s ds}$, which represents the present value of receiving one dollar at time $t$. Since this is a price of a risk-free zero coupon bond maturing at time $t$, I use the cubic spline approach to bootstrap the zero-coupon bond prices for the necessary maturities from the yield curve of constant maturity treasury prices (Longstaff, Mithal, and Neis (2005)). For the LIBOR/SOFR rate, $q_t$, I assume that the one month rate is held fixed to the one month rate at the time of the CRT tranche issuance, since the first coupon payment is known at time of issuance since coupon payments tied to reference rates are set in advance. Variation in the reference rate could affect the yield on investments in CRT bonds, but since defaults and prepayments are modeled independent from interest rates in this model it is not a crucial piece of the exercise.

## 2.4 Empirical Motivation of the Two Factor Model

Before estimating the model in the following section, I provide further motivation for such a specification using historical data on mortgage defaults from the Fannie Mae single-family data set. The two factor model is motivated principally by the fact that mortgage market participants and commentators often consider credit risk through the lens of routine default behavior and risks of catastrophic loss, such as those experienced during the Subprime Mortgage Crisis. This qualitative two factor framework can be seen, for example, in Goodman et al. (2014) where the authors consider the setting of g-fees both in a normal and stress scenario. In this paper, I formalize the two factor model and estimate mortgage losses in those types of scenarios. This is naturally modeled in the above framework as two Poisson default intensities, one common and one rare, that trigger low and high default rates respectively.

I leverage the Fannie Mae Single-Family data set in order to investigate the historical default behavior during two periods that represent these two notions of default risk. Loans originated during the early parts of the single-family dataset, particularly 1999-2001, experienced low to normal levels of default and are often considered a good baseline for mortgage default rates in benign market conditions. On the opposite end of the spectrum are loans originated between 2005 and 2007, which experienced the brunt of the housing downturn and the Global Financial Crisis. In the FNMA dataset, I focus on the default rates, loss severities and time between entering default and loan disposition as the primary objects of interest. This is in line with the parameters that govern defaults in the credit model described before.

Recall that the parameter vector governing default dynamics in the model will be specified by $\Theta \in [\gamma_1, \gamma_2, lgd_1, lgd_2, del_1, del_2]$. In this parameter vector, $\gamma_i$ represents the amount of notional principal that enters default upon realization of a given default event $i \in [1, 2]$. $lgd_i$ represents the loss given default in a particular default scenario, also called the loss severity. Since mortgage loans are secured by the underlying loan collateral, it is highly unlikely that default results in the total loss of outstanding principal on the loan. An analogous object would be the recovery rate in the modeling of corporate bond pricing. Lastly, the length of the liquidation process, $del_i$, is another relevant modeling object and is thus included in the estimation exercise; mortgage foreclosure proceedings can be long and expensive, and lenders often do not realize losses on a particular loan until months after payments have ceased being made.

### 2.4.1 Parameters: $\mathbb{P}$ vs. $\mathbb{Q}$

Instead of estimating the full set of risk-neutral parameters from the data by optimizing the model to match CRT tranche prices, I match the parameters to similar empirical moments from the single-family data set. This is equivalent to making the assumption that the observed parameter vector is equal the risk-neutral one. Here $\Theta^{\mathbb{Q}} = \Theta^{\mathbb{P}}$, and only the state-variables $\lambda_1$ and $\lambda_2$ differ between the risk-neutral and physical probability measures. It is possible that the market expects different size jump sizes, for example. It is likely that this assumption, however, is relatively innocuous as differing jump sizes would change the associated risk-neutral intensities and potentially wash out when comparing implied loss distributions. Directly specifying parameters as gleaned from historical data also have the added benefit of allowing one to compare how the market views the probability of events that compare to those that have shaken the market in the past.

Estimating the full set of risk-neutral parameters from tranche prices could provide powerful insights for future research and for the estimation of risk premiums pertaining to housing finance. For example, markets may view severities in crisis scenarios as higher than experienced historically, or write-downs occurring sooner after loans enter default. While we cannot measure the physical probabilities of such events in this framework, this would provide suggestive evidence that their are risk premiums embedded along other dimensions of mortgage credit risk that have yet to be discovered.

With this caveat in mind, I now discuss the motivation of the calibrated parameters below, with an emphasis on how their values may differ during both normal and crisis times. This provides further support for the particular functional form of the two factor default model as well as justification for why the model provides a good fit.[15]

15. Anecdotal evidence, of course, is highly suggestive that a quasi two factor structure is present in how investors price CRT trances. For example, Belbase (2014) values early CRT tranches using a variety of scenarios which weigh crisis events with base case scenarios.

### 2.4.2 Jump Sizes

The jump size parameter, $\gamma_i$, is the proportion of underlying loan notional principal balance that enters default at the realization of a single default event driven by $\lambda_i$. Recall that $\gamma_1$ is interpretable as a baseline level of default; even during normal times, when the housing market is healthy and house prices are stable, borrowers may default for a variety of reasons. To choose $\gamma_1$, I observe baseline default rates in the FNMA data for the benign cohort years 1999-2001. These default rates are defined as the fraction of loans in a cohort that reach 180 days delinquent in a given month.[16]

Appendix table 2.5 shows that these baseline monthly default rates are between 0.75 and 0.96 basis points. For ease of exposition, I choose $\gamma_1$ to be equal to 1 basis point, which has the nice interpretation of being equal to 10 loans defaulting in a typical CRT cohort size of 100,000 loans. Results using values between 0.75 and 1 basis points have little effect because the estimated state variables will just adjust to be slightly higher to match the appropriate loss level.

Appendix table 2.6 shows the maximum yearly default rates incurred in each of the three crisis year cohorts as expressed in a rolling 12 month window. For example, the worst 12 months that the 2007 cohort experienced was a year in which 4.14% of loans reached 180 days delinquent. Why would we be interested in total delinquency rates over extended periods rather than during just one month? This has to do with limiting the structure that we have to place on the problem of determining how defaults cluster across time periods in a crisis scenario. Since the model considers independent Poisson processes, the Poisson arrival of the second risk factor could be

16. I acknowledge that this is not exactly one to one with the concept of default in the model, where upon realization of a default event loans immediately stop paying and do not ever become current again. I choose 180 days delinquent to calibrate the parameters because it is typically given as the standard definition of mortgage "default." The difference is likely to be small, and in fact one could argue that any small rounding errors made due to discrepancies in this definition could be washed out in the severity rate.

interpreted as causing a cluster of defaults that do not necessarily have to occur in the same month in reality even though they do in the model. Longstaff and Rajan (2008) note a similar interpretation, which in their paper takes the form of a rare event striking the corporate bond sector. Thus simply looking at the highest one month default rates will not be revealing as defaults strongly cluster in subsequent months across a crisis. This intuition is likely supported in the data, where the chosen $\gamma_2$ of 0.032 results in a strong fit to the CRT tranche prices. I choose this value as roughly the average of the three crisis cohorts considered in the data.

### 2.4.3 Loan Loss Severities

Analogous to recovery rates on corporate bonds, mortgage loans have a loss severity that is simply 1 less the recovery rate. Table 2.7 shows the average loss severities for loans in each of the cohorts studied in this section. Severities are markedly higher during the crisis cohorts. The mean loss severity during the 1999-2001 cohorts varies from 36 to 46%, whereas during the crisis cohorts it ranged from 51-55%. The 1999-2001 cohorts have a much lower median severity, due to the nature of the distribution of severities during that time which I discuss next.

Figure 2.10 shows smoothed densities for the loans in those cohorts upon which default occurred and a subsequent loss was recorded. The first striking thing is the difference in the shape of the probability density function: the crisis years feature a much more pronounced hump shape where as the benign cohort years are strongly right skewed.[17] This is consistent with several other studies on the severities associated with mortgage default, as well as the different default behavior of loans depending on the economic context of the situation; Goodman and Zhu (2015) provides a historical account of mortgage loss severities on GSE loans and An and Cordell (2021)

17. Severities are not necessarily bounded, as the costs associated with the foreclosure process can be on the same order of magnitude as the remaining principal on the loan when it defaults.

provides context on post-crisis loss severities.

Given this data, I choose $[lgd_1, lgd_2] = [.33, .55]$ to optimize the model; the two factor structure of mortgage defaults is well supported in the data when it comes to loss severities. There is a markedly different point estimate, as well as distribution, of severities in normal times versus crisis times.

### 2.4.4 Time Between Default and Loss Realization

Another relevant question to both the management of mortgage credit risk and the design of CRT bond programs is the time between when a loan enters default and when a loss is subsequently realized. To do so I look at the length between the last time a payment was made and the disposition date. Why is this important for CRT investors? Until the loss is written down, bondholders will continue to receive interest on the current amount of tranche principal balance.

Table 2.8 shows summary statistics for liquidation times (in months) for both each cohort year as well as terciles sorted by eventual loss severity. Median liquidation times range from 18 to 26 months. There is a clear relation between the time that a loan spends in disposition and the eventual severity on the loan. Figure 2.11 below plots the kernel density of the time between entering default and realizing a loss. Observations are grouped into terciles of eventual severity. I find that loans with eventual high loss severities also take the longest to resolve - an intuitive finding, but one that has very interesting implications for credit risk transfer. Since the most senior CRT bonds are paid down first, defaults would have to be extremely bad in a crisis scenario in order to write-down the most senior CRT bonds. O'Neill (2022) argues that the probability of such a scenario in the lens of the market is not zero, otherwise the value of the senior tranches would be equal to their risk free rate.

Given this data, I choose $[del_1, del_2] = [18, 24]$ when fitting the model. The data on default rates, severities, and liquidation times are strongly suggestive of the two factor model being a good candidate for fitting implied defaults from CRT tranche prices.

## 2.5 Estimation and Results

The model was fitted as described above by discretizing the paydown formulas and utilizing the Monte Carlo pricing method and the Subplex optimization algorithm as described in the following section to fit the levels of the state variables. Table 2.1 below shows the parameter vector I use for the exercise, as motivated in the previous section using historical data on conforming GSE loans.

| $\gamma_1$ | $\gamma_2$ | $del_1$ | $del_2$ | $lgd_1$ | $lgd_2$ |
|---|---|---|---|---|---|
| 1e-04 | 0.032 | 18 | 24 | 0.33 | 0.55 |

**Table 2.1:** The parameters used in the model estimation. $\gamma_1$ and $\gamma_2$ represent the jump sizes of the first and second default processes.

Market prices are taken from the TRACE data set for the first week of trading of each particular tranche, identified by its CUSIP from the deal-level prospectus data. To estimate the model, the tranche cash flow formulas are discretized and it is assumed that cash flows occur at month end. Thus the estimated intensities can be interpreted as the arrival of a default event during month $t$ rather than instantaneously. Appendix section 2.A.9 shows examples of how the formulas are converted into the discrete-time approach used in O'Neill (2022).

### 2.5.1 Simulation Methodology

Due to the lack of a closed-form solution for the valuation of each CRT tranche, I use Quasi-Monte Carlo simulation in order to solve for the risk neutral state variables $\lambda_{1t}$ and $\lambda_{2t}$ for each bond issuance in the sample. Along each sample path, I simulate the realization of the two Poisson processes. Conditional on the realization of each default process, the valuation process collapses to simply evaluating each expression for the cash flow waterfalls of each particular tranche. Monthly cash flows and write downs occur at the end of each month in the discretized model. We can consider the realizations of each Poisson process as a $T \times 2$ random matrix $X$, where rows represent each month until bond maturity and columns represent the two state variables. This results in effectively over 700 random variables (since $T > 350$ for the issuances), in which their order matters due to the paydown structure of the bonds.

The valuation procedure as described can be reduced to the following form below in equation 2.10, where $PV_x(K_j, \Delta_i, \Theta, X)$ is the present value function for one realization of the default states for tranche with attachment points $K_j$ as given in equation 2.17. $\Delta_i$ is a vector of deal specific information such as time to maturity of the bonds, tenor of the CRT tranches, the underlying mortgage interest rate, and the discount curve $D(t)$ on the valuation date.

$$PV(K_j, T) = \int PV_x(K_j, \Delta_i, \Theta, X) p(X) dX = \frac{1}{N} \sum_{k=1}^{N} PV_x(X_k, ...) \tag{2.10}$$

$N$ is the number of sample paths, or simulations of $X$ in the Monte Carlo procedure. Standard Monte Carlo techniques can evaluate this integral but at prohibitively large computational cost. Confounding the simulation problem is the well-known issue associated with simulating "rare events." I find that values of $\lambda_{2t}$ imply a realization around once every 2-300 months, but it is important to be aware that realizations early in the life of the pool are the most consequential for the valuation of the bonds.

Due to the high dimensional nature of the random variables, I opt for Quasi-Monte Carlo (QMC) methods. Using the Sobol Sequence generation of Joe and Kuo (2008), convergence for the Monte-Carlo pricing algorithm is achieved and I am able to move on to the optimization of the state variables.

### 2.5.2 Numerical Optimization

The optimization step involves looping over the pricing simulation until convergence is achieved by finding the level of the state variables that have the lowest pricing errors for a given set of parameters $\Theta$. I define the pricing error as the root mean squared error (RMSE) of tranche market prices relative to the model implied prices; each deal issuance therefore has their own set of state variables but share a common parameter vector $\Theta$. For a deal $i$ with $J$ CRT tranches and market prices MP, the non-linear least squares optimization problem is stated as follows:

$$\min_{\lambda_{i,1},\lambda_{i,2}} \sqrt{\frac{\sum_{j=1}^{J}(PV(K_j) - MP(K_j))^2}{J}} \tag{2.11}$$

For optimizing over the state space, I utilize the Subplex method of Rowan (1990), which can be considered a variant of a Nelder-Mead algorithm that can accept box constraints.[18] The function to be optimized is well behaved in the sense that extreme state levels of the state variables in either direction quickly make the bonds either risk-free or worthless, so the general region of solution is identified. However, since I have no formal proof of convexity and the function is non-linear, I opt for the Subplex method to ensure efficient searching of the entire state space (due to periodic restarts) in order to be conservative and ensure that the global minimum has been reached.

### 2.5.3 Estimation Results

Recall that $\gamma_1$ and $\gamma_2$ represent the jump sizes, which can be interpreted as the portion of the outstanding notional balance that enter into default in a given period.

18. Constraints on $\lambda_i$ were set to $[0, 0.1]$ and $[0, 10]$ for $\lambda's$ 1 and 2, respectively. Because of the nature of the Subplex method, which introduces periodic restarts to the Nelder-Mead method, results are robust to choice of starting values. Upper and lower bounds are wide enough to capture both no credit risk, and a level of credit risk that makes the bonds effectively worthless.

Therefore, a realization of the state variable $\lambda_2$ would result in 3.2% of loans in the pool entering default. After a time period specified by $del_2$, losses are incurred on the balance that entered default would be equal to the defaulted balance multiplied by the the severity (loss given default), $lgd_2$.

Figure 2.2 below shows the time-series and cross-section of root mean squared errors from fitting the model, expressed in basis points per $100 notional. The model is fit for each bond issuance so that there is one set of state variables for each issuance that best prices the 4+ tranches sold in that particular issuance. Results are grouped based on the LTV group of the bond issuance. The RMSE for this particular parameter vector is 97 basis points, and often lower. There are several outliers, particularly with group 2 issuance, suggesting that the market may be viewing a set of parameters $\Theta$ that are further from the set I have used in this paper, particularly during the COVID-19 crisis when we see one deal RMSE reach almost 300 basis points.

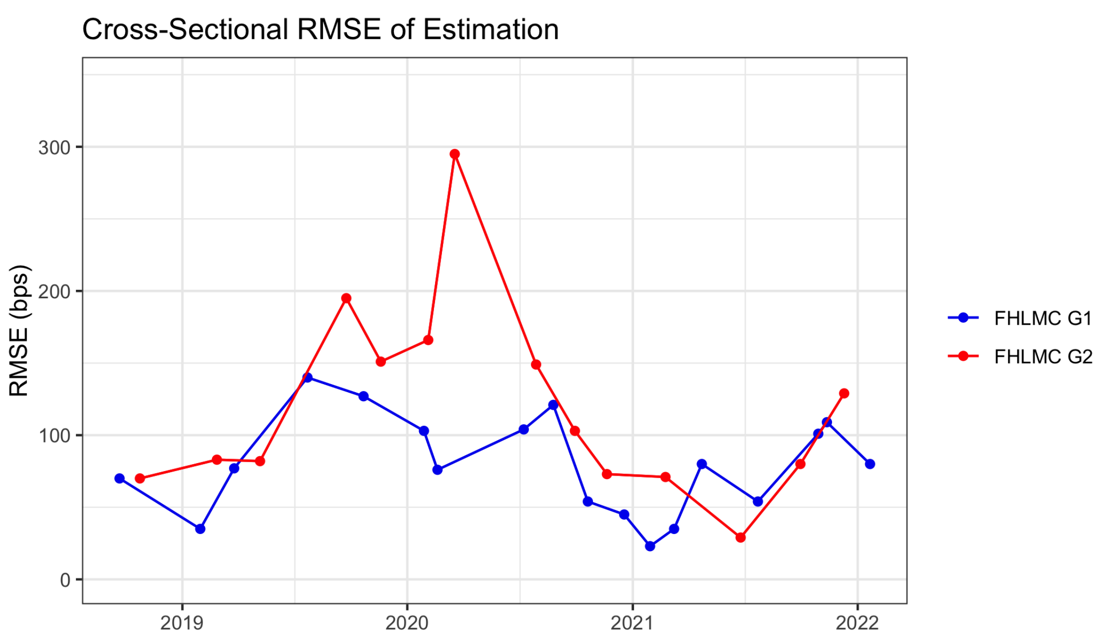



**Figure 2.2:** This figure plots the time-series root-mean-square-error of fitting the valuation model to tranche prices for each GSE group-level issuance. The RMSE is measured in basis points per $100 bond notional.

Despite the outliers discussed above, I argue that this model represents a good fit given the relative simplicity of the model. Future research could involve both specifying an alternate parameter vector for Group 2 deals or fully estimating a risk-neutral set of parameters $\Theta$, as long as the restriction that the model was still over-identified holds. Table 2.2 shows summary statistics for deal level state variables and tranche RMSEs.

| | $\lambda_1$ | $\lambda_2$ | $RMSE(bps)$ |
|---|---|---|---|
| Mean | 1.897 | 0.0035 | 97.188 |
| SD | 0.48 | 0.00147 | 54.727 |
| Min | 1.003 | 0.00155 | 23 |
| Max | 2.722 | 0.00653 | 295 |
| N | 32 | 32 | 32 |

**Table 2.2:** Estimation Results

The estimated state variables are shown in figure 2.3. Panel A shows the time series of the first state variable, $\lambda_1$. Recall that $\lambda_1$ can be thought of as the routine or expected credit risk of the mortgage pool. From the plot, we see that there is still time variation in the risk-neutral estimates of $\lambda_1$, with the majority of the intensities for various bond issuance falling around 2. An intensity of 1, with an estimated $\gamma_1$ of 0.0001 translates into the expectation that 10 loans default in a given month in a diversified pool of 100,000 mortgages. This is interesting and suggests that the market may require a substantial risk premium for holding onto mortgage credit risk that is generally considered "predictable" in the sense that econometric models can forecast the baseline levels of default under the $\mathbb{P}$ measure with a considerable degree of accuracy.

Panel B shows the time series of the second state variable, $\lambda_2$; $\lambda_2$ is picking up the catastrophic risk associated with the mortgage pool. We see that this state variable exhibits substantial time series variation, peaking in 2020 when market risk premiums were reaching highs due to the COVID-19 pandemic. Again, as a back of the envelope

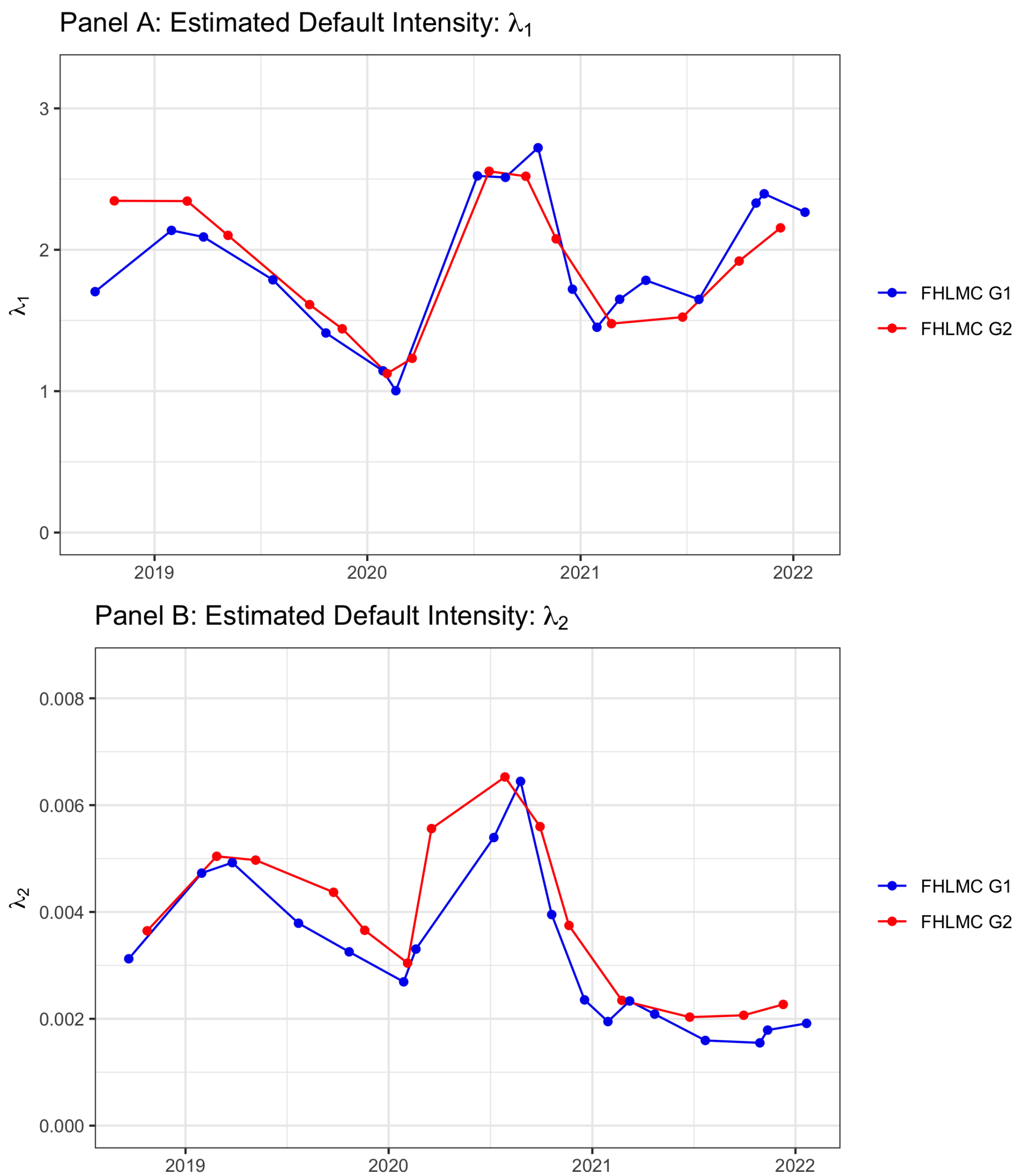



**Figure 2.3:** This figure plots the time-series of estimated default state variables. Panel A plots $\lambda_1$, the state variable corresponding to individual mortgage specific default risk; panel B plots, $\lambda_2$, the catastrophic default risk. See text for discussion and a description on interpreting the magnitudes in the context of mortgage pool losses.

calculation in an unseasoned pool, an intensity of 0.0035 represents a credit event that occurs roughly every 285 months (24 years). The implications of this credit event, at the estimated jump size $\gamma_2$ of 0.032 are substantial; roughly 3200 mortgages in a given month. This level of crisis can be considered on the level of the 2008 mortgage crisis, and this model provides direct evidence that investors in CRT bonds price this risk and are concerned about a repeat mortgage crisis despite improved underwriting standards post-crisis.

Appendix 2.A.8 shows results broken down by LTV group. There were 18 Group 1 deals included in the estimation and 14 Group 2 deals. Group 1 deals have a lower RMSE by almost 50 basis points; table 2.9 also shows that Group 1 deals have a slightly lower level of catastrophic risk on average. Furthermore, average levels, as well as the distribution of $\lambda_1$ are very similar across both groups. This could have the interpretation that the market views private mortgage insurance (PMI) as being relatively fail safe in normal times but could break down under a catastrophic scenario.

### 2.5.4 How do we assess the model errors?

One question is whether the errors are "small," since we lack the traditional asset pricing context on what constitutes a good level of estimation error since I am pricing securities and not returns. One way to assess whether the two factor model offers an improvement is by assessing the performance of estimating only a single-factor model. Appendix section 2.A.13 shows the results for attempting to estimate the model with only a single factor. Three paramaterizations of the single factor model are considered. Figures 2.14 and 2.15 consider the same parameters as both $\lambda_1$ and $\lambda_2$ in the main model, but estimated separately as their one factor counterpart. Unsurprisingly, estimating the model with $\lambda_1$ as the only risk factor leads to extremely high pricing errors, with an RMSE of about 1300 basis points on average. Intuitively, pricing errors follow the same pattern as floater spreads; high spreads on mezzanine tranches lead to high risk free values, and since the small jump size in this model can not cause defaults to strike the mezzanine tranches, their pricing errors track the spreads.

Moving to Figures 2.15, the $\lambda_2$ factor alone fares better, but is still drastically worse than the two factor model, with a mean RMSE of 660 basis points. Lastly, 2.15 bridges the gap between the two models and estimates a factor with parameters that are the average of those used in the baseline specification. Again, this model performs much worse than the two factor model but does surprisingly well in the latter stages of the estimation period. This is consistent with the evidence in section 2.6.2, where I show that the contribution of the implied g-fee is increasingly driven by one factor, or that the risk of catastrophic default has fallen in recent times. Thus a model ignoring this component is able to do relatively better, but not as well as the two factor model. This section provides further evidence that the market believes defaults are driven by more than one factor, although it does not establish that the two factor model is the globally optimal model.

## 2.6 Simulation Study

The remainder of the paper takes the view that we can learn about the market view of mortgage risk through the risk factors estimated above. In order to do so, I simulate many paths of the pricing model using the parameters at their starting values and risk factors constant at their optimized values. I then pull out objects of interest from the pricing model in order to evaluate the market view of mortgage credit risk; for example, losses caused by type 2 defaults or the weighted average life of the pools. This study takes the implicit assumption that the estimated state variables provide a consistent view of the market price of risk in the mortgage pool and that defaults can continue to happen in the same way even after all the CRT tranches are paid off;

that is, residual risk or risk not covered by the CRT tranches is subject to the same no-arbitrage loss distribution.

The primary objects that I look at are as follows: firstly, the distribution of losses implied by the levels of the state variables. These loss distributions are analogous to those calculated in O'Neill (2022), and we can directly compare the beta implied loss distributions to the loss distributions calculated via simulation here as well as the implied g-fees over all sample paths. More importantly, the principal contribution of this paper is that the richer model can allow us to make inferences that O'Neill (2022) could not. For example, I break down the implied g-fee into its components stemming from both risk factors and find that both risk factors make roughly equal contributions to the implied g-fee, with the routine level of default risk playing an increasing role in recent issuances. Next, I analyze the credit coverage provided by GSE bonds. I discuss both the assumptions required to make such inferences and the limits to what CRT bonds can tell us. I do find that at least suggestively, CRT bonds provide credit coverage in most scenarios.

### 2.6.1 Simulation Implied Loss Distributions and G-fees

As noted above, I use the model parameters and optimized state variables to simulate the paydown of the underlying mortgage pools and observe. The output of this model are the risk-neutral collateral dynamics implied by the CRT tranche prices. Figure 2.4 shows the simulation of the loss distributions for a sample of four different bond issuances. The outputs appear bimodal, which is expected as the two state variables will produce discrete loss scenarios. Much lower levels of tail risk produce very interesting dynamics in the bottom two-plots. Note that STACR 2020 HQA3 was one of the bond issuances during the summer of 2020 during the COVID-19 pandemic, and as a result it appears that risk premiums were elevated. The second hump shape occurs around the point of $\gamma_2$ multiplied by $lgd_2$, shifted to the right by the amount that is on average the losses due to the first risk factor.

Furthermore, one can calculate a a market implied g-fee, that is the fair insurance premium that one would have to pay in order to insure against default losses. The credit cost portion of the g-fee is comprised of expected losses and a risk-premium component. Since the losses implied by the pricing model are estimated under the risk-neutral measure, the expectation of the implied loss distribution is equal to this g-fee credit cost after we have amortized it over the life of the loan, as shown below:

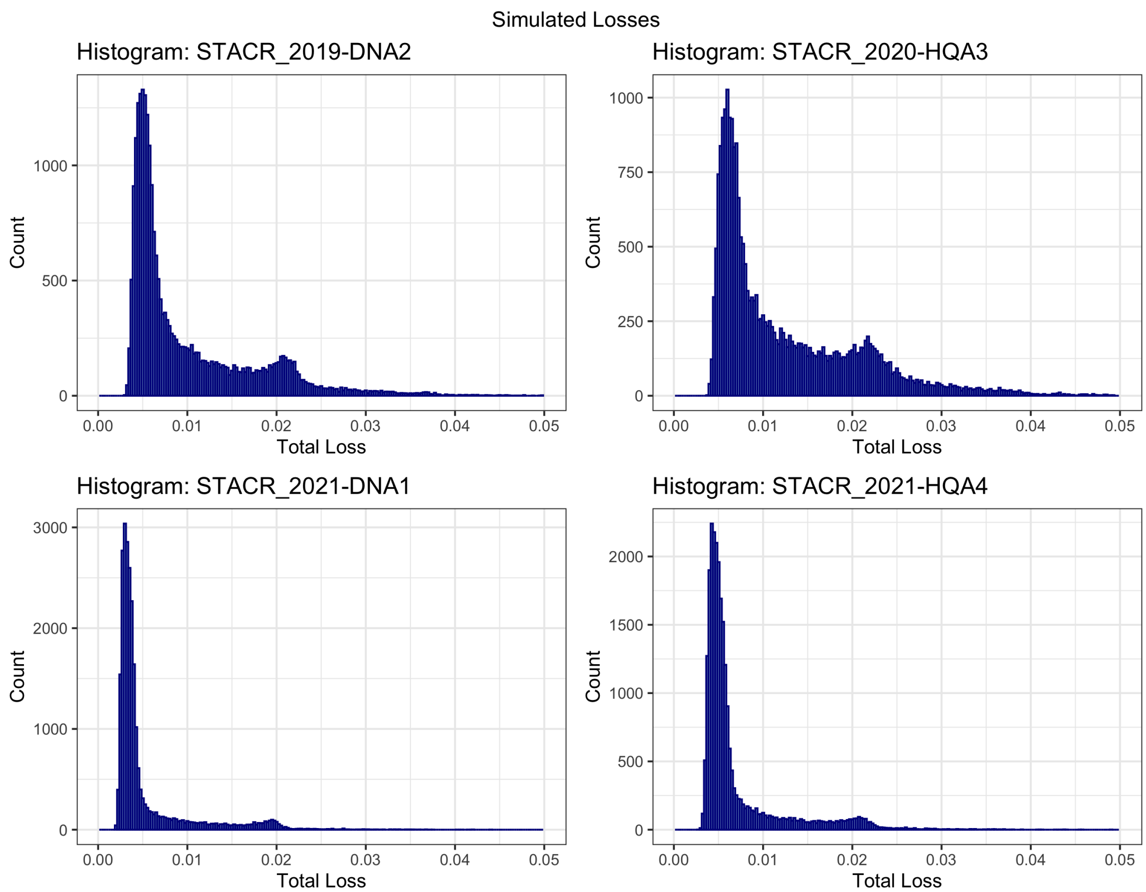



**Figure 2.4:** Histogram from simulating defaults under optimized state variables for several select bond issuances. The two factor model naturally produces loss distributions that are bimodal.

$$\text{g-fee} = \mathbb{E}^{\mathbb{P}}[CL/WAL] + RiskPremium = \underbrace{\mathbb{E}^{\mathbb{Q}}[CL/WAL]}_{\substack{\text{Identified by} \\ \text{CRT prices}}} \tag{2.12}$$

The g-fee can be calculated along each simulation path, along with a path specific weighted average life (WAL). Thus each simulation represents a fair g-fee in each scenario, meaning that the model can speak to the whole distribution of the g-fees rather than simply the point estimate. Implied g-fees are shown in figure 2.5. The g-fee ranges from 10 to 20 basis points, confirming the results in O'Neill (2022) that market-implied g-fees are lower than those currently charged by the GSEs, which are around 35 basis points. Some further references on the g-fee and its relationship to CRT prices can be found in Palmer (2017), Goodman et al. (2014), deRitis and Zandi (2014), Richardson, Van Nieuwerburgh, and White (2017), Elenev, Landvoigt, and Van Nieuwerburgh (2016), Belbase (2014), and O'Neill (2022).

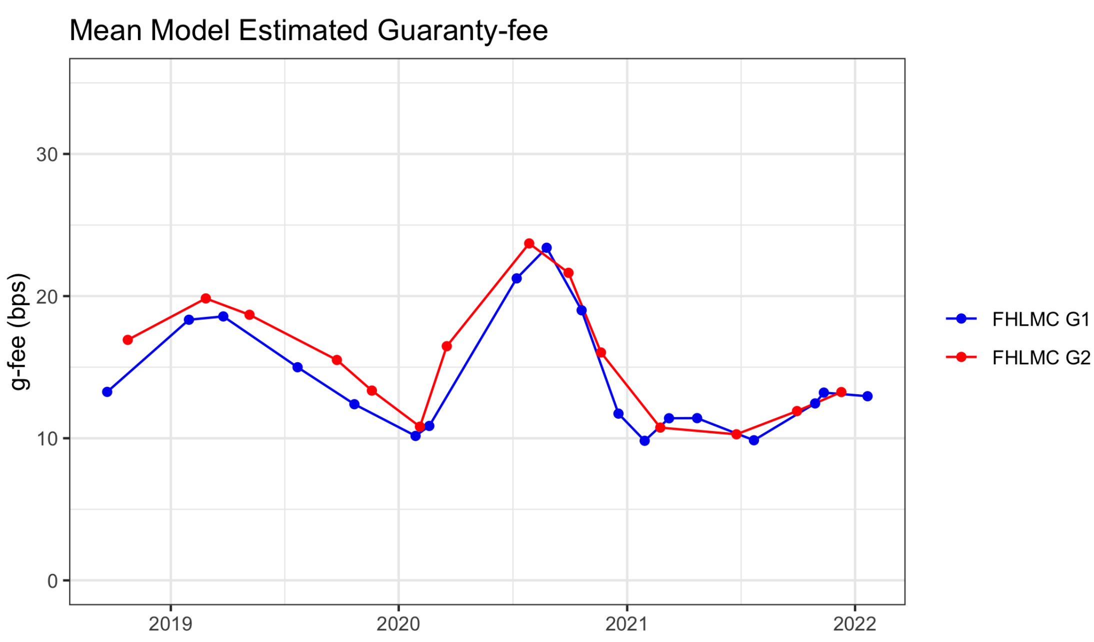



**Figure 2.5:** Time-series plot of implied g-fees by bond issuance. The comparison between $\lambda_2$ and the g-fee shows the extent to which the g-fee is determined by the risk premium for the catastrophic default risk.

The full distribution of g-fees is described in appendix section 2.A.11; these distributions could be valuable to the GSEs from a VaR type perspective. For example, the GSEs may wish to set g-fees in order to be compensated fairly in the worst 75% of scenarios, rather than just in the mean case. In the following section, 2.6.2, I provide a further breakdown of the GSE g-fee and the particular facets of it that our model can speak to.

### 2.6.2 Contribution to Guaranty-Fee

One analysis that the two factor model allows is to decompose the g-fee into contributions from both types of mortgage risk. This is similar to Bhansali, Gingrich, and Longstaff (2008), who demonstrate that spreads in the CDO index market are well described by a model that effective decomposes CDO spreads into firm-level, sectoral, and global risk and that the relative proportions of these factor in CDO tranche spreads vary over time. Figure 2.6 below performs this analysis, where the implied g-fee for a given bond issuance is decomposed into expected losses from both risk factors. While both risk factors have been contributing roughly equally over time, expected losses due to $\lambda_1$ have slowly increased since 2020.

Since this plot is normalized as a proportion, it may have the interpretation that the probability of defaults have increased as of late but that the market does not view catastrophic risk as increasing. The arrival of new CRT issuances and slowing house price appreciation as of the writing of this article could bring new insights from this model, as we may again see catastrophic risk rise.

#### 2.6.2.1 Comparison to the O'Neill (2022) Model

O'Neill (2022) uses tranche market prices to fit an implied loss probability density function. Above, I compute a similar density function by simulating the model using

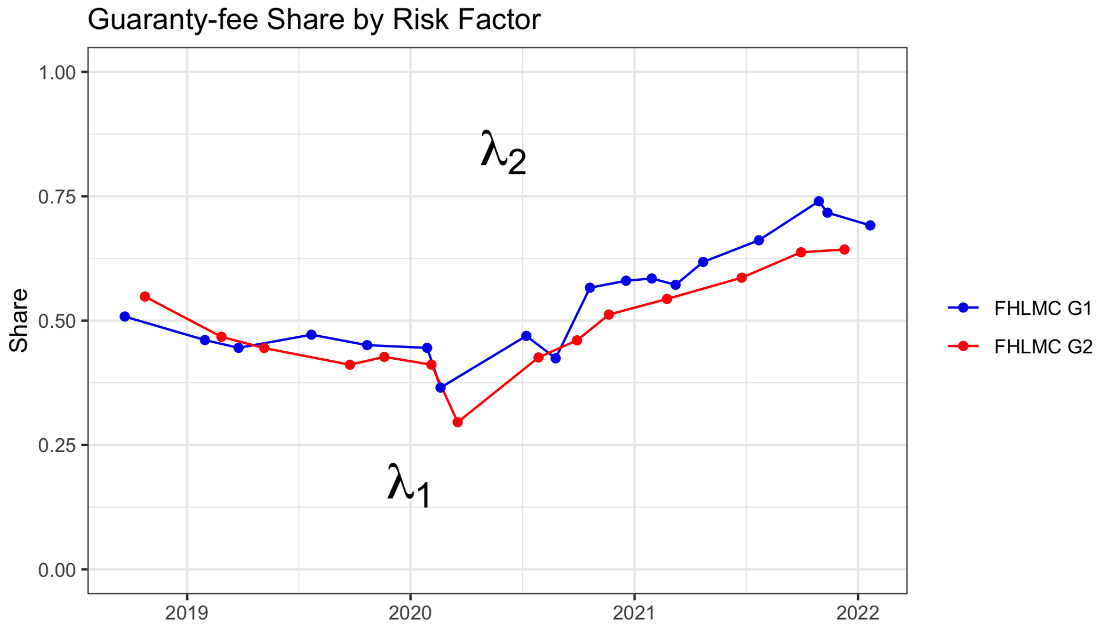



**Figure 2.6:** This plot shows contributions to the g-fee from each risk factor calculated by simulating the pool losses due to each type of risk.

the optimized state variable levels and plotting a histogram of the results. A natural follow up question is how these histograms compare to the fitted beta distributions from earlier work. Appendix section 2.A.10 plots two example issuances along with the implied loss distribution in red from O'Neill (2022). The distributions differ in a few interesting ways. First of all, the right tail of the distribution is generally approximated well, but the nature of the two factor model precludes low levels of losses between 0 and 25 basis points. An improvement on the technique in O'Neill (2022) could be to censor the distribution around 25 basis points, which would in spirit be a restriction that the risk neutral distribution of losses is strictly greater than the expectation of the physical one, or that there is no scenario in which the market places a positive probability on there being less defaults than the econometrically estimated level of losses.

### 2.6.3 Credit Coverage

Lastly, this section uses the simulation approach to evaluate the design of the CRT bonds. Several outstanding questions related to CRT design are well-adapted to my model, which characterizes the distribution of losses implied by tranche market prices. The most important of these questions is how well CRT bonds provide credit protection to the issuing GSEs. In the simulation study I calculate the total losses incurred on the collateral pool as well as the total write-downs attributed to CRT bonds. The fraction of loans covered by the bonds is the "credit coverage" of the deal issuance implied by the prices of CRT bonds. Figure 2.7 below shows the mean credit coverage for each deal issuance along with the coverage in the worst 1, 5 and 10% of loss scenarios.

As mentioned above, this could be used in a value-at-risk type analysis where the GSEs use tranche market prices to assess how the market views the potential of loan losses to exceed the protection afforded to them by issuing the CRT bonds. For example, in the 5th percentile of worst loss scenarios, the GSEs retain between 10 and 20% of the residual credit risk. The GSEs will have to trade off these risks against the cost of issuing the bonds to arrive at the conclusion of whether "credit risk transfers transfer credit risk" Quantitatively, this question is ill-posed without a model, because without the risk-neutral distribution of credit coverage, econometric specifications of mortgage default may make the right tail appear much less probable than markets expect it to be.

This analyses is subject to a few caveats. For example, I am implicitly assuming that the estimated dynamics can be extrapolated to default behavior on the loans which will not be covered by the CRT bonds. One can think of a plausible scenario in which the loans which do not prepay under any circumstance also exhibit unique default behavior. Despite this, I believe that my approach is a good first step towards unifying qualitative studies of CRT bonds with a more rigorous asset pricing

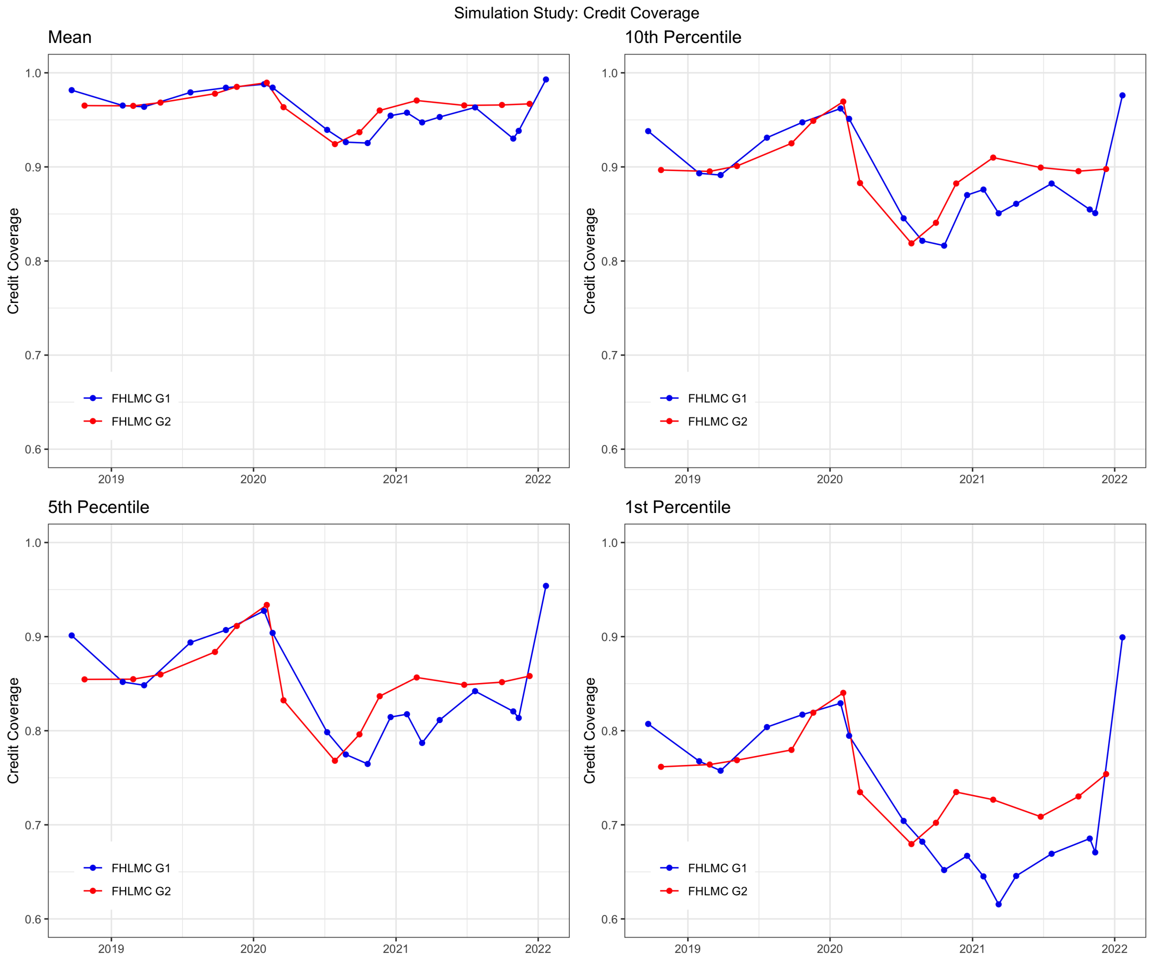



**Figure 2.7:** Simulated Credit Coverage of STACR Issuances.

treatment. Future research may be able to disentangle these effects more cleanly.

## 2.7 Conclusion

In just under nine years, the Credit Risk Transfer market has quietly transformed the landscape of the US mortgage market and transferred the credit risk on trillions of dollars of mortgages to private investors. The program is subject to intense public debate; it's proponents argue that the CRT programs mean that the taxpayer is now protected, where as their detractors argue the structure of CRT deals obfuscate the true level of credit protection they are providing. This paper contributes to that debate by being one of the first academic papers to study this unique market, and to my knowledge, the first to formally value tranche prices in a reduced-form credit model.

This paper provides a top-down two factor no-arbitrage model for the valuation of CRT bond tranches. In the model, the two factors represent both routine and catastrophic mortgage risk. The model fits tranche prices well, with a tranche level root mean squared error of 97 basis points, and leads to a number of interesting implications. First and foremost, the parameters implied by the model allow us to directly calculate the mortgage credit risk premium on the mortgage pools underlying the bond issuances and therefore a market-based g-fee.

The model can also speak to the relative contributions of both types of credit risk that are priced in CRT tranches. It can provide a live snapshot of the market base probability of a disaster event in the mortgage market. It can also speak to how the market views the level of credit protection in the mortgage market. I find that even at detachment points around 4% for most senior CRT tranche, the market expects the GSE's to be on the hook for losses about 5% of the time.

There are numerous questions that further research should seek to address. For example, matching prices and spreads in the secondary market to create daily measures of the forward looking mortgage risk premium. A future model could also estimate the full parameter vector of risk-neutral parameters and potentially uncover risk-premiums among other dimensions of credit risk. An econometric model for expected defaults under the $\mathbb{P}$ measure could also further identify the exact risk premiums on the factors, although such a model is not necessary for getting the g-fee since it includes a risk premium component. I hope that the modeling techniques introduced here serve as a helpful starting point for such studies.

## 2.A Appendix

### 2.A.1 Sample Summary Statistics

| Deal | Issued | Group | Tranches | Attach | Detach | Tenor | WAC | WAM | WCS | WLTV |
|---|---|---|---|---|---|---|---|---|---|---|
| STACR_2018-DNA3 | 2018-09-21 | 1 | 4 | 0.10 | 4.00 | 360 | 4.34 | 354.00 | 741.97 | 75.80 |
| STACR_2018-HQA2 | 2018-10-24 | 2 | 4 | 0.10 | 4.00 | 360 | 4.19 | 351.00 | 741.05 | 92.60 |
| STACR_2019-DNA1 | 2019-01-30 | 1 | 4 | 0.10 | 4.25 | 360 | 4.76 | 353.00 | 739.62 | 75.97 |
| STACR_2019-HQA1 | 2019-02-26 | 2 | 4 | 0.10 | 4.50 | 360 | 4.67 | 353.00 | 737.49 | 92.93 |
| STACR_2019-DNA2 | 2019-03-26 | 1 | 4 | 0.10 | 4.25 | 360 | 4.92 | 354.00 | 740.10 | 76.21 |
| STACR_2019-HQA2 | 2019-05-07 | 2 | 4 | 0.10 | 4.50 | 360 | 4.82 | 352.00 | 736.87 | 93.02 |
| STACR_2019-DNA3 | 2019-07-23 | 1 | 4 | 0.10 | 4.25 | 360 | 4.96 | 351.00 | 741.30 | 76.00 |
| STACR_2019-HQA3 | 2019-09-24 | 2 | 4 | 0.10 | 4.50 | 360 | 4.92 | 351.00 | 736.70 | 92.91 |
| STACR_2019-DNA4 | 2019-10-22 | 1 | 4 | 0.10 | 4.00 | 360 | 4.98 | 350.00 | 740.15 | 75.91 |
| STACR_2019-HQA4 | 2019-11-19 | 2 | 4 | 0.10 | 4.50 | 360 | 4.98 | 352.00 | 736.73 | 92.75 |
| STACR_2020-DNA1 | 2020-01-28 | 1 | 4 | 0.10 | 3.75 | 360 | 4.60 | 350.00 | 745.05 | 76.04 |
| STACR_2020-HQA1 | 2020-02-04 | 2 | 4 | 0.10 | 4.25 | 360 | 4.51 | 354.00 | 741.89 | 92.80 |
| STACR_2020-DNA2 | 2020-02-19 | 1 | 4 | 0.10 | 3.75 | 360 | 4.19 | 352.00 | 750.76 | 76.13 |
| STACR_2020-HQA2 | 2020-03-18 | 2 | 4 | 0.10 | 4.00 | 360 | 4.12 | 354.00 | 745.91 | 92.58 |
| STACR_2020-DNA3 | 2020-07-08 | 1 | 4 | 0.25 | 4.00 | 360 | 3.94 | 352.00 | 753.31 | 75.51 |
| STACR_2020-HQA3 | 2020-07-28 | 2 | 4 | 0.25 | 4.00 | 360 | 3.85 | 352.00 | 750.77 | 91.97 |
| STACR_2020-DNA4 | 2020-08-25 | 1 | 4 | 0.25 | 4.00 | 360 | 3.95 | 353.00 | 752.39 | 75.45 |
| STACR_2020-HQA4 | 2020-09-29 | 2 | 4 | 0.25 | 4.00 | 360 | 3.86 | 353.00 | 750.39 | 92.01 |
| STACR_2020-DNA5 | 2020-10-20 | 1 | 4 | 0.10 | 3.50 | 360 | 3.57 | 354.00 | 758.12 | 74.79 |
| STACR_2020-HQA5 | 2020-11-20 | 2 | 4 | 0.25 | 3.75 | 360 | 3.45 | 355.00 | 752.76 | 91.18 |
| STACR_2020-DNA6 | 2020-12-18 | 1 | 4 | 0.25 | 3.00 | 360 | 3.40 | 354.00 | 758.23 | 74.37 |
| STACR_2021-DNA1 | 2021-01-29 | 1 | 4 | 0.25 | 2.50 | 360 | 3.26 | 353.00 | 762.10 | 74.45 |
| STACR_2021-HQA1 | 2021-02-23 | 2 | 4 | 0.25 | 3.25 | 150 | 3.16 | 355.00 | 755.13 | 90.88 |
| STACR_2021-DNA2 | 2021-03-09 | 1 | 4 | 0.25 | 2.50 | 150 | 3.09 | 354.00 | 762.53 | 74.30 |
| STACR_2021-DNA3 | 2021-04-23 | 1 | 4 | 0.25 | 2.50 | 150 | 2.97 | 353.00 | 763.40 | 74.15 |
| STACR_2021-HQA2 | 2021-06-25 | 2 | 4 | 0.25 | 3.00 | 150 | 2.92 | 353.00 | 755.94 | 90.67 |
| STACR_2021-DNA5 | 2021-07-23 | 1 | 4 | 0.25 | 2.00 | 150 | 2.94 | 352.00 | 762.80 | 73.88 |
| STACR_2021-HQA3 | 2021-09-30 | 2 | 4 | 0.25 | 3.25 | 240 | 2.81 | 353.00 | 753.80 | 90.37 |
| STACR_2021-DNA6 | 2021-10-29 | 1 | 4 | 0.25 | 2.00 | 240 | 2.86 | 351.00 | 759.91 | 73.88 |
| STACR_2021-DNA7 | 2021-11-12 | 1 | 4 | 0.25 | 2.25 | 240 | 3.03 | 354.00 | 755.17 | 74.28 |
| STACR_2021-HQA4 | 2021-12-10 | 2 | 4 | 0.25 | 3.50 | 240 | 3.00 | 354.00 | 752.55 | 91.07 |
| STACR_2022-DNA1 | 2022-01-21 | 1 | 5 | 0.25 | 4.50 | 240 | 3.15 | 355.00 | 751.98 | 74.92 |

**Table 2.3:** CRT Deal Summary Statistics. This table contains the subset of CRT deals for which the reduced-form credit model is estimated.

### 2.A.2 Sample Prices

| Deal | M1 | M2 | B1 | B2 |
|---|---|---|---|---|
| STACR_2018-DNA3 | 100.18 | 100.51 | 101.54 | 99.85 |
| STACR_2018-HQA2 | 100.01 | 99.88 | 99.94 | 103.35 |
| STACR_2019-DNA1 | 100.13 | 100.65 | 101.51 | 100.98 |
| STACR_2019-DNA2 | 100.02 | 100.09 | 100.31 | 101.17 |
| STACR_2019-DNA3 | 100.09 | 100.15 | 101.35 | 102.80 |
| STACR_2019-DNA4 | 100.00 | 100.00 | 100.13 | 101.04 |
| STACR_2019-HQA1 | 100.08 | 100.50 | 100.50 | 100.76 |
| STACR_2019-HQA2 | 100.10 | 100.21 | 100.60 | 106.21 |
| STACR_2019-HQA3 | 100.02 | 99.99 | 100.06 | 102.37 |
| STACR_2019-HQA4 | 100.04 | 100.08 | 100.12 | 100.00 |
| STACR_2020-DNA1 | 100.04 | 99.98 | 100.35 | 103.44 |
| STACR_2020-DNA2 | 100.09 | 100.12 | 98.93 | 98.86 |
| STACR_2020-DNA3 | 100.25 | 100.19 | 100.62 | 100.56 |
| STACR_2020-DNA4 | 100.14 | 100.52 | 101.60 | 101.08 |
| STACR_2020-DNA5 | 100.19 | 100.12 | 101.07 | 103.67 |
| STACR_2020-DNA6 | 100.05 | 100.02 | 100.22 | 101.33 |
| STACR_2020-HQA1 | 100.06 | 100.34 | 99.86 | 101.45 |
| STACR_2020-HQA2 | 97.86 | 97.12 | 95.17 | 100.19 |
| STACR_2020-HQA3 | 100.08 | 99.53 | 99.65 | 99.80 |
| STACR_2020-HQA4 | 100.12 | 100.39 | 100.80 | 100.36 |
| STACR_2020-HQA5 | 100.00 | 100.23 | 100.38 | 102.94 |
| STACR_2021-DNA1 | 100.09 | 99.97 | 100.14 | 100.92 |
| STACR_2021-DNA2 | 100.06 | 100.12 | 99.72 | 100.06 |
| STACR_2021-DNA3 | 100.07 | 100.86 | 100.64 | 100.78 |
| STACR_2021-DNA5 | 100.00 | 100.51 | 101.56 | 101.76 |
| STACR_2021-DNA6 | 100.04 | 100.16 | 100.24 | 100.96 |
| STACR_2021-DNA7 | 100.02 | 100.08 | 100.19 | 101.33 |
| STACR_2021-HQA1 | 100.04 | 99.89 | 100.00 | 99.19 |
| STACR_2021-HQA2 | 100.08 | 100.13 | 100.60 | 103.70 |
| STACR_2021-HQA3 | 100.03 | 100.19 | 100.23 | 100.73 |
| STACR_2021-HQA4 | 99.99 | 100.11 | 100.36 | 100.92 |
| STACR_2022-DNA1 | 100.10 | 100.10 | 99.90 | 98.98 |

**Table 2.4:** CRT Deal Market Prices. This table contains the market prices of the subset of CRT deals for which the reduced-form credit model is estimated.

### 2.A.3 Attachment/Detachment Points

The below plot shows the attachment and detachment points for the 32 CRT bond issuances used to estimate the reduced-form credit model. I limit the sample to include recent issuances with 4 tranche prices and junior bonds with attachment points equal or lower than 25 basis points to aid in the identification of $\lambda_1$. There is still substantial heterogeneity in the offered tranches as can be seen from the plot below.

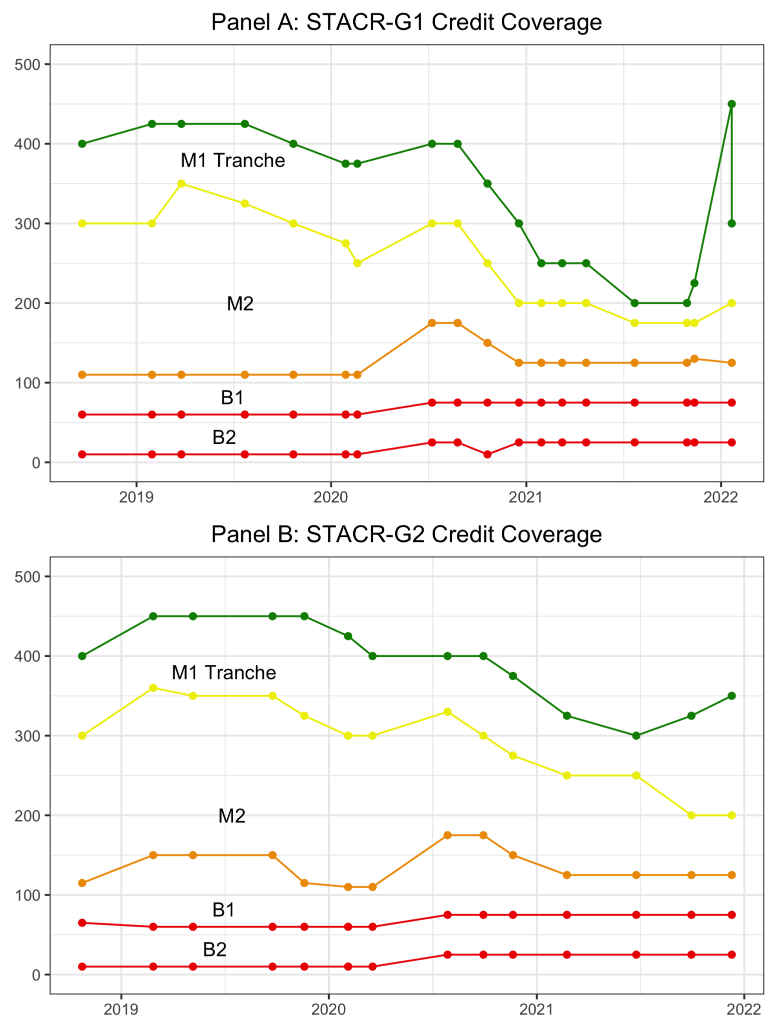



**Figure 2.8:** This plot shows the credit coverage for the subset of CRT bonds estimated using the model in this paper.

### 2.A.4 Sample Floater Spreads

The below plot shows the floater spreads for the 32 CRT bond issuances used to estimate the reduced-form credit model. Spreads peaked in 2020 during the COVID-19 pandemic, but since tranche attachment and detachment points vary over time, spreads for a generic tranche name are not necessarily directly comparable over time without a model.

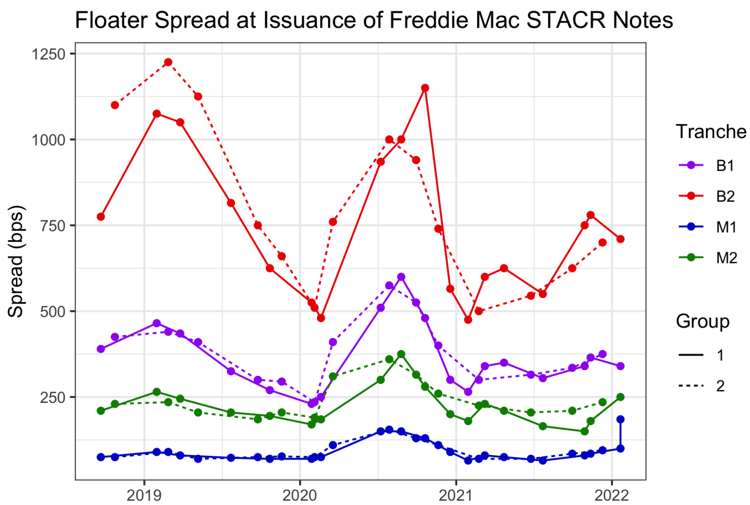



**Figure 2.9:** This plot shows the floater spread for each tranche of the subset of CRT deals estimated using the model in this paper.

### 2.A.5 Jump Sizes

| Cohort | n Loans | Mean | Median | p25 | p75 |
|---|---|---|---|---|---|
| 1999 | 158065 | 0.00009620 | 0.00005061 | 0.00001898 | 0.00011388 |
| 2000 | 1266734 | 0.00008037 | 0.00004263 | 0.00001362 | 0.00009000 |
| 2001 | 2799525 | 0.00007642 | 0.00005822 | 0.00001661 | 0.00011931 |

**Table 2.5:** Monthly cohort default rates expressed as a percentage of the total loans in the cohort

| Cohort | n Loans | Max |
|---|---|---|
| 2005 | 1445682 | 0.0216 |
| 2006 | 1079542 | 0.0329 |
| 2007 | 1111526 | 0.0414 |

**Table 2.6:** Cumulative cohort default rates expressed as the highest 12 month cumulative default percentage in each of the crisis year cohorts.

### 2.A.6 Loss Severities

| Cohort | n Loans | Prin ($b) | n Loss Rlzd | Total loss ($b) | Mean | Median | p25 | p75 |
|---|---|---|---|---|---|---|---|---|
| 1999 | 158065 | 18.9 | 1416 | 0.030 | 36.25 | 19.62 | 3.47 | 54.63 |
| 2000 | 1266734 | 160.6 | 12368 | 0.294 | 40.74 | 23.46 | 3.61 | 64.29 |
| 2001 | 2799525 | 391.3 | 26031 | 0.817 | 46.77 | 36.07 | 8.06 | 74.17 |
| 2005 | 1445682 | 252.2 | 79388 | 6.141 | 51.66 | 48.66 | 28.60 | 70.17 |
| 2006 | 1079542 | 198.4 | 81662 | 7.508 | 55.52 | 53.53 | 33.45 | 74.54 |
| 2007 | 1111526 | 216.2 | 90045 | 8.189 | 52.80 | 50.19 | 29.70 | 72.45 |

**Table 2.7:** Loss Severity Summary Statistics

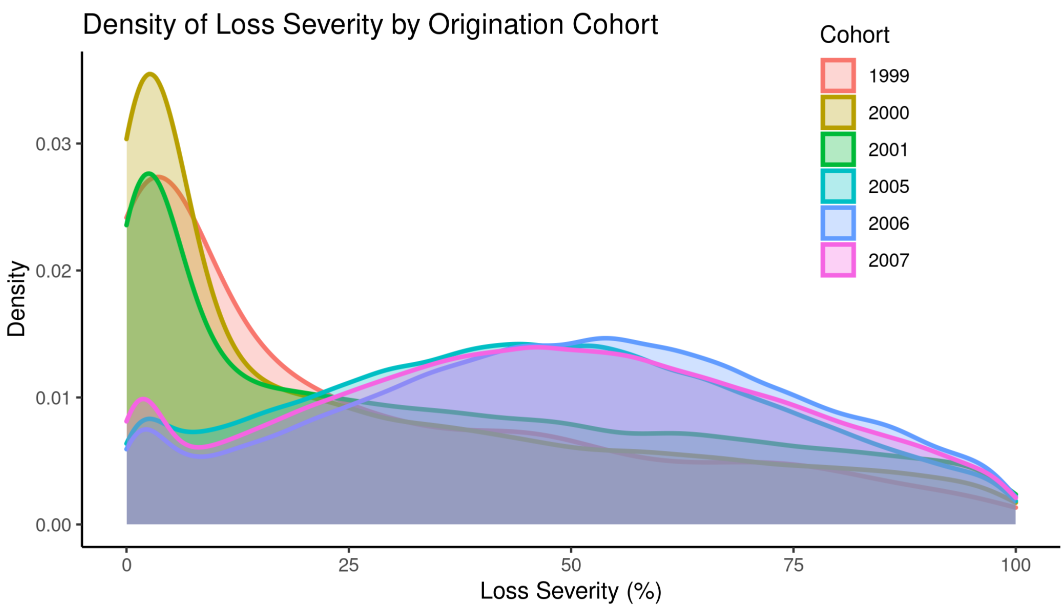



**Figure 2.10:** Distribution of Loss Severities in the FNMA Historical Data set

### 2.A.7 Time to Liquidation

| Cohort | n Loss Rlzd | Mean | Median | p25 | p75 |
|---|---|---|---|---|---|
| 1999 | 1416 | 21.3 | 18 | 13 | 25 |
| 2000 | 12368 | 21.5 | 18 | 14 | 25 |
| 2001 | 26031 | 22.4 | 19 | 14 | 26 |
| 2005 | 79388 | 24.8 | 20 | 13 | 31 |
| 2006 | 81662 | 24.8 | 20 | 13 | 31 |
| 2007 | 90045 | 25.1 | 20 | 13 | 32 |
| High | 96970 | 32.7 | 26 | 18 | 42 |
| Low | 96971 | 18.5 | 16 | 11 | 23 |
| Mid | 96969 | 22.5 | 19 | 13 | 28 |

**Table 2.8:** Summary Statistics: Time between default and write-down for loans that experienced losses.

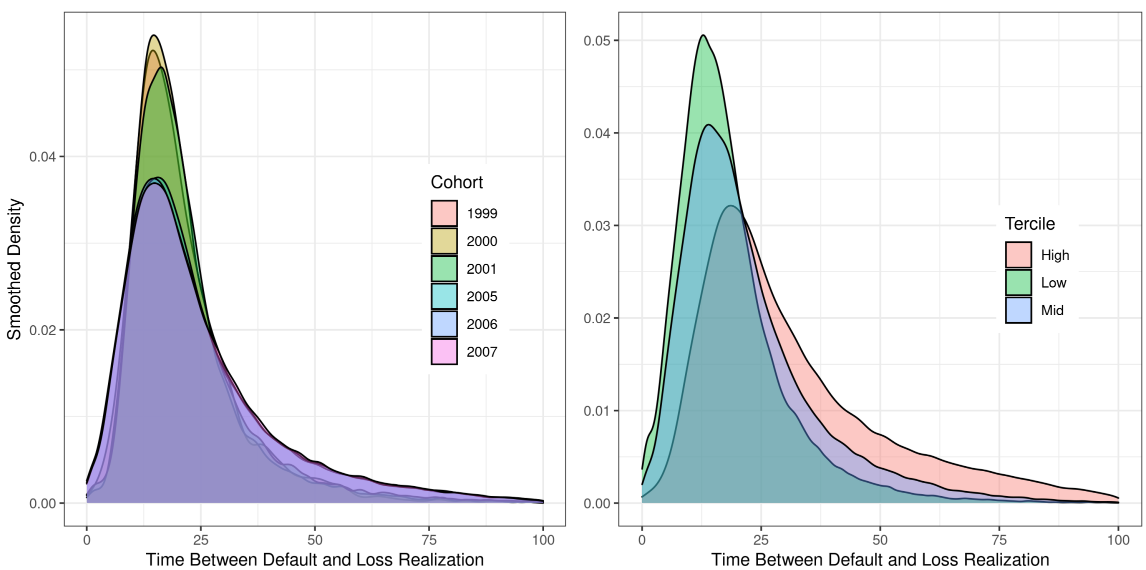



**Figure 2.11:** Time Between Defaults and Loss Realization

### 2.A.8 Estimation Results by LTV Group

| | $\lambda_1$ | $\lambda_2$ | $RMSE(bps)$ |
|---|---|---|---|
| Mean | 1.905 | 0.00318 | 79.667 |
| SD | 0.493 | 0.00144 | 34.193 |
| Min | 1.003 | 0.00155 | 23 |
| Max | 2.722 | 0.00645 | 140 |
| N | 18 | 18 | 18 |

**Table 2.9:** Estimation Results for Group 1

| | $\lambda_1$ | $\lambda_2$ | $RMSE(bps)$ |
|---|---|---|---|
| Mean | 1.888 | 0.00392 | 119.714 |
| SD | 0.481 | 0.00147 | 68.129 |
| Min | 1.125 | 0.00203 | 29 |
| Max | 2.555 | 0.00653 | 295 |
| N | 14 | 14 | 14 |

**Table 2.10:** Estimation Results for Group 2

### 2.A.9 Discrete Mortgage Cash Flows

This section shows the mortgage cash flows when the model is converted to discrete time. Many of the formulas are standard in the calculation of the paydown of mortgage-backed securities. I have adopted several of them to account for the defaults featured in my model, which are not featured in credit guaranteed MBS. For example rather than being continuously compounded, the total notional not yet prepaid would be expressed as:

$$P_t = \prod_{n=1}^{t} 1 - p_n \tag{2.13}$$

Cash flows are assumed to only occur at the end of a given month. Scheduled principal and interest, and prepayments come from the standard mortgage amortization formulas. As shown in O'Neill (2022), calculating $B_t^*$ is equivalent, as is the calculation of $Q_t$ as the product of $D_t$ and $P_t$. Rather than having an instantaneous default rate, any default that occurs during a given month is recognized at the end of the monthly period. Thus the default rate at time $t$ would be:

$$d_t = \gamma_1 \Delta N_{1,t-1\to t} + \gamma_2 \Delta N_{2,t-1\to t} \tag{2.14}$$

And the amount of defaulted balance:

$$def_{i,t} = B_t^* \times d_t \tag{2.15}$$

Lastly, written down losses at the end of period $t$ due to default risk $i$ would be calculated as:

$$l_{i,t} = B_t^* \times lgd_i \times \gamma_i \Delta N_{i,t-1\to t} \tag{2.16}$$

The tranche present value equation for a tranche with detachment point $K_j$ and maturity $T$ becomes:

$$PV(K_j, T) = \sum_{t=1}^{T-1} D(t) \times E_t^Q[\underbrace{B_{K_j,t-1}(q_t + s_{K_j})}_{\text{Coupon Payment}} + \underbrace{PC_t - PC_{t-1}}_{\text{Principal Repayment}}] + \underbrace{D(T)B_{K_j,T}}_{\text{Principal at Maturity}} \tag{2.17}$$

### 2.A.10 Implied Loss Distributions Comparison

Figure 2.12 compares simulated implied loss distributions with the fitted Beta distributions for the same CRT issuances as calculated in O'Neill (2022), which are shown overlaid in red.

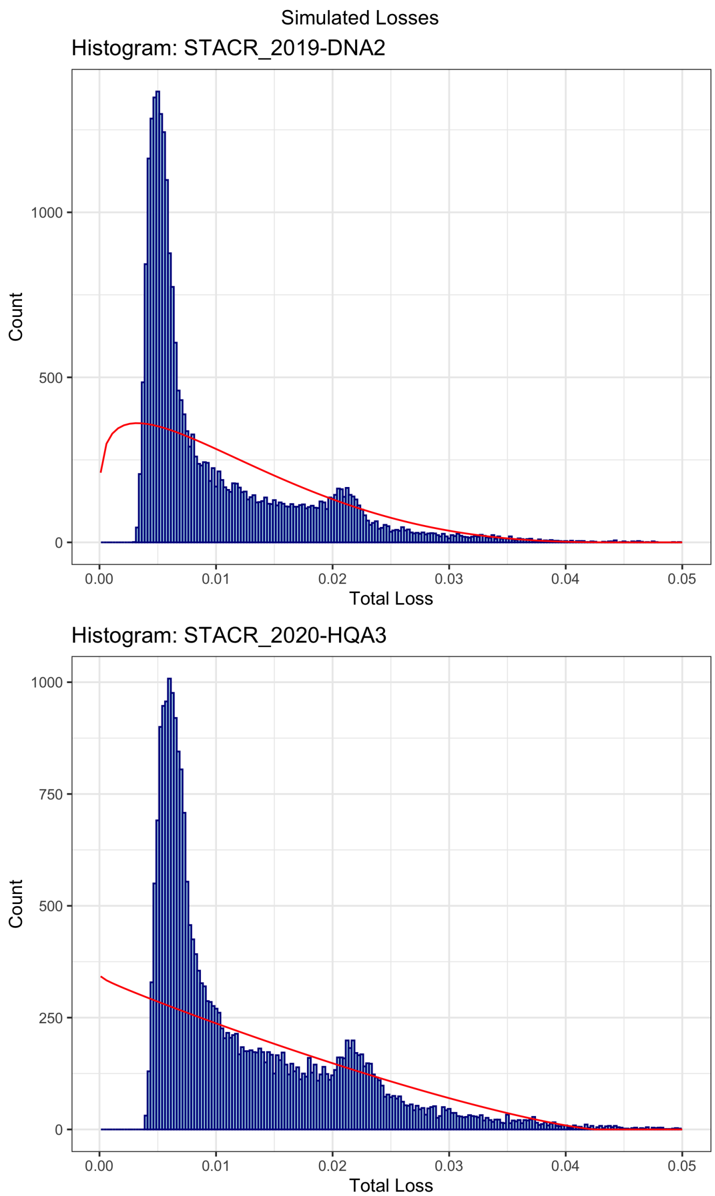



**Figure 2.12:** caption place holder

### 2.A.11 Implied g-fee Distributions

The analysis in the main section of the paper focuses on the implied g-fee as being the expectation of the g-fee under the estimated paydown parameters. The total distribution of g-fees is of empirical interest too from a VaR-type perspective. The following plot shows the time series of estimated g-fees at the 50th, 75th, and 95th percentiles in addition to the mean values already presented in the paper.

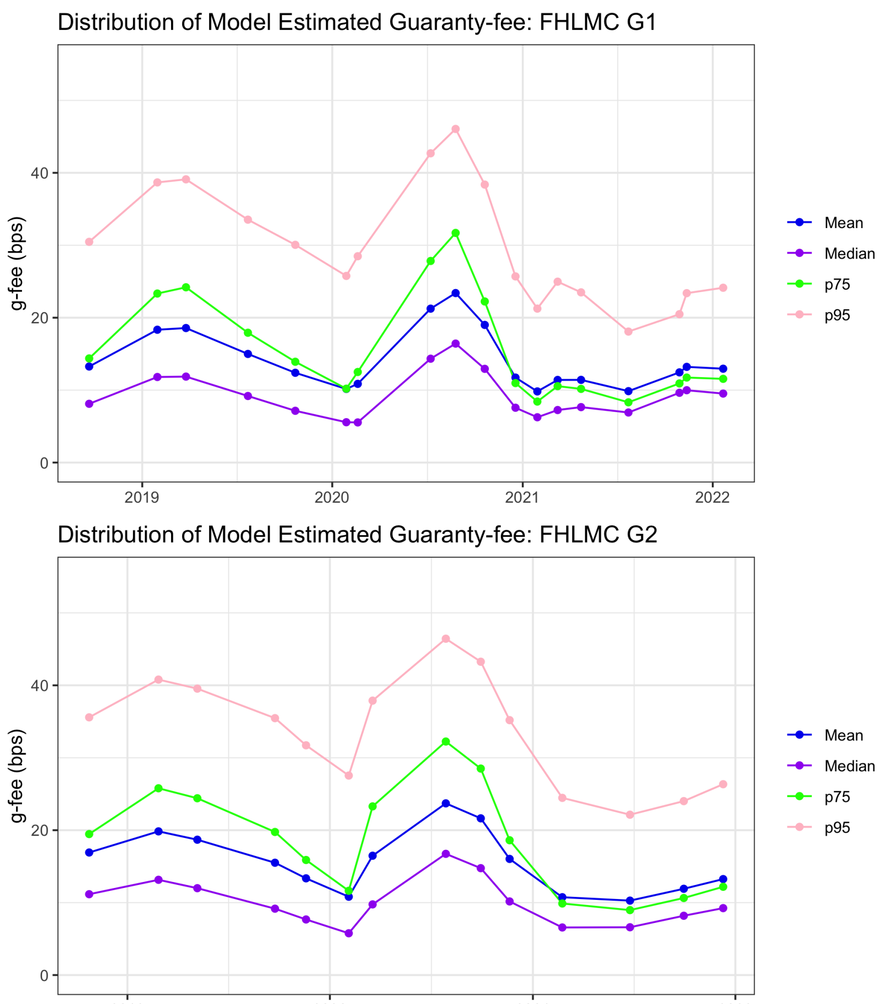



**Figure 2.13:** Time-series plots of implied g-fees by bond issuances. These plots show the mean, median, and 75th/95th percentiles of the g-fee.

### 2.A.12 Other Features of CRT Bonds

There are several added layers of complexity involved in CRT deals that are worth mentioning. Most notably are two different triggers that effect the timing of cash flows to the bond investors. The first is referred to as the minimum enhancement trigger; the CRT reference pools must maintain a certain level of credit enhancement to the senior retained tranche (A) in order for the subordinated classes to receive unscheduled principal payments. Often, this level of credit enhancement is not met upon the issuance of a bond. For example, a deal may have the minimum enhancement test is 4.75%. At issuance, credit enhancement stood at only 4%. Thus the subordinated classes will not receive any unscheduled principal, or prepayments, until the senior class has been paid down enough to meet this requirement. This is a potentially important, but albeit complex, feature of these bonds, and future iterations of this project would benefit from their inclusion. At this time, the pricing model does not account for them.

The second is called the delinquency trigger. In the event that there is significant distress in the underlying mortgage pool, subordinated classes will be cut off of receiving prepayments. This is done in order to ensure that the senior classes and the GSE are protected from losses that could accrue to them if the junior classes were paid off too quickly. The exact triggers have varied over time and by issuance. For example, in the CAS program, if 40% of the subordinate principal balance is 90+ days delinquent, unscheduled principal is cut off to the subordinated classes. This is likely less important in the model, as defaults are modeled as a one-stage event upon which they never become current. Furthermore, crisis events can typically wipe out the entire subordinated classes meaning they would not receive any more principal anyways.

Lastly, some issuance of CRT bonds have so-called "recombinable" notes, which offer investors the chance to exchange their notes in pre-specified amounts for notes with different coupon payments. The market for these recombination is small, illiquid, and in many cases, no one has yet exercised the option to recombine at the time of writing of this paper. The economic value of this option is likely to be small, and is likely to add little value to the analysis presented here.

### 2.A.13 Single-factor Model

This section presents results from estimating several versions of a single factor model and showing that they are unable to price the CRT tranches nearly as accurately as the two factor model advocated in this paper. Specifications (A) and (B) simply use the same parameters as each state-variable estimated in the two factor model. Specification (C) uses parameters as the average of the two specifications.

Results from One-Factor Estimation

| $\gamma$ | del | lgd |
|---|---|---|
| 1e-04 | 18 | 0.33 |

| | $\lambda$ | RMSE(bps) |
|---|---|---|
| *Mean* | 2.48 | 1294.767 |
| *SD* | 0.473 | 532.505 |
| *Min* | 1.75 | 486 |
| *Max* | 3.438 | 2136 |
| *N* | 30 | 30 |

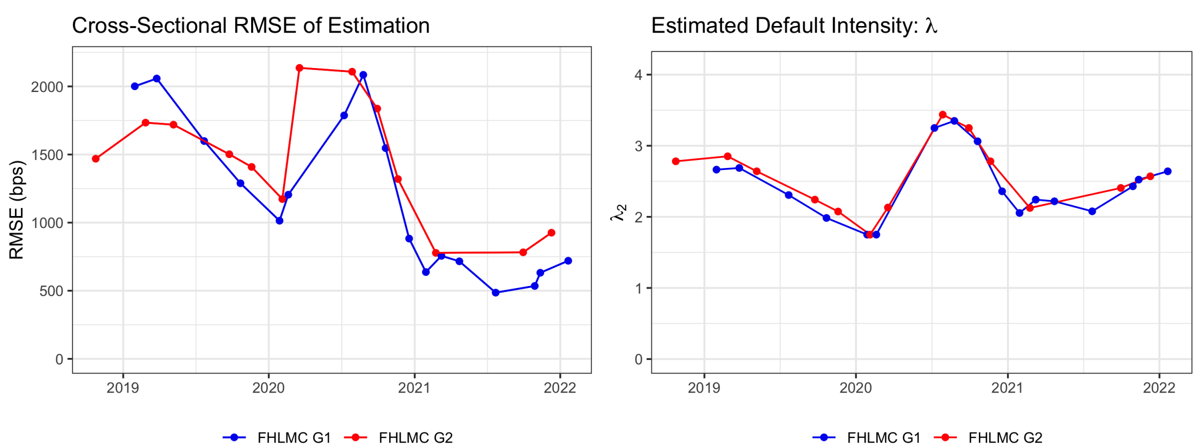



**Figure 2.14:** One Factor Specification (A) Results: Clockwise, the parameters used in the estimation, the results, the estimated default intensity, and the cross-sectional tranche RMSE for each CRT bond deal.

Results from One-Factor Estimation

| γ | del | lgd |
|---|---|---|
| 0.032 | 24 | 0.55 |

| | λ | RMSE(bps) |
|---|---|---|
| *Mean* | 0.007 | 660.719 |
| *SD* | 0.003 | 310.008 |
| *Min* | 0.003 | 252 |
| *Max* | 0.012 | 1350 |
| *N* | 32 | 32 |

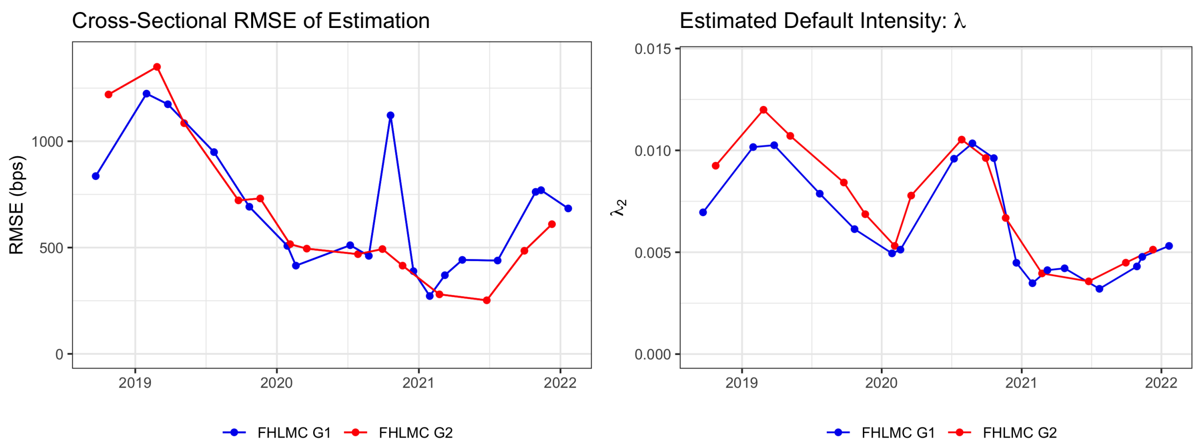



**Figure 2.15:** One Factor Specification (B) Results: Clockwise, the parameters used in the estimation, the results, the estimated default intensity, and the cross-sectional tranche RMSE for each CRT bond deal.

Results from One-Factor Estimation

| γ | del | lgd |
|---|---|---|
| 0.01605 | 21 | 0.44 |

| | λ | RMSE(bps) |
|---|---|---|
| *Mean* | 0.013 | 370.531 |
| *SD* | 0.004 | 193.264 |
| *Min* | 0.008 | 15 |
| *Max* | 0.021 | 830 |
| *N* | 32 | 32 |

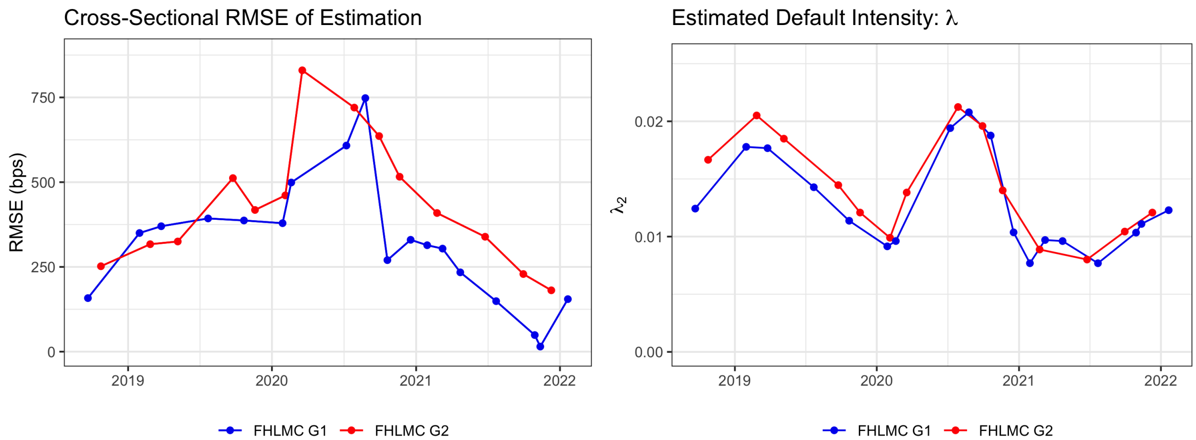



**Figure 2.16:** One Factor Specification (C) Results: Clockwise, the parameters used in the estimation, the results, the estimated default intensity, and the cross-sectional tranche RMSE for each CRT bond deal.

## References

An, Xudong, and Larry Cordell. 2021. "Mortgage loss severities: What keeps them so high?" Real Estate Economics 49 (3): 809–842.

Belbase, Eknath. 2014. "STACR and CAS: What They Tell Us." The Journal of Structured Finance 19 (4): 45–51. https://doi.org/10.3905/jsf.2014.19.4.045.

Bhansali, Vineer, Robert Gingrich, and Francis A. Longstaff. 2008. "Systemic Credit Risk: What Is the Market Telling Us?" Financial Analysts Journal 64 (4): 16–24.

Bluhm, Christian, and Christoph Wagner. 2011. "Valuation and Risk Management of Collateralized Debt Obligations and Related Securities." Annual Review of Financial Economics 3 (1): 193–222.

Chernov, Mikhail, Brett R. Dunn, and Francis A. Longstaff. 2017. "Macroeconomic-Driven Prepayment Risk and the Valuation of Mortgage-Backed Securities." The Review of Financial Studies 31, no. 3 (December): 1132–1183.

deRitis, Cristian, and Mark Zandi. 2014. "A General Theory of G-Fees." Economic and Consumer Credit Analytics.

Diener, Nicolas, Robert Jarrow, and Philip Protter. 2012. "Relating Top-Down with Bottom-Up approached in the Evaluation of ABS with Large Collateral Pools." International Journal of Theoretical and Applied Finance 15 (02): 1250011.

Duffie, Darrell, and Nicolae Garleanu. 2001. "Risk and Valuation of Collateralized Debt Obligations." Financial Analysts Journal 57 (1): 41–59.

Duffie, Darrell, and Kenneth J. Singleton. 1999. "Modeling Term Structures of Defaultable Bonds." The Review of Financial Studies 12 (4): 687–720. Accessed April 26, 2022.

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh. 2016. "Phasing out the GSEs." Journal of Monetary Economics 81 (C): 111–132.

Fermanian, Jean-David. 2013. "A Top-Down Approach for Asset-Backed Securities: A Consistent Way of Managing Prepayment, Default and Interest Rate Risks." The Journal of Real Estate Finance and Economics 46, no. 3 (April): 480–515.

Finkelstein, David, Andreas Strzodka, and James Vickrey. 2018. "Credit Risk Transfer and De Facto GSE Reform." Economic Policy Review 24 (3): 557–595.

Gao, Chao, and John J. McConnell. 2018. "Investment Performance of Credit Risk Transfer Securities (CRTs): The Early Evidence." The Journal of Fixed Income 28 (2): 6–15. ISSN: 1059-8596.

Gauthier, Laurent. 2003. "Market-Implied Losses and Non-Agency Subordinated MBS." The Journal of Fixed Income 13 (1): 49–74. ISSN: 1059-8596. https://doi.org/10.3905/jfi.2003.319346. eprint: https://jfi.pm-research.com/content/13/1/49.full.pdf.

Giesecke, Kay, Lisa R. Goldberg, and Xiaowei Ding. 2011. "A Top-Down Approach to Multiname Credit." Operations Research 59 (2): 283–300.

Golding, Edward, and Deborah Lucas. 2020. "Credit Risk Transfer and the Pricing of Mortgage Default Risk." MIT Golub Center Working Paper.

Goodman, Laurie, Ellen Siedman, Jim Parrott, and Jun Zhu. 2014. "Guarantee Fees - An Art, Not a Science." Housing Finance Policy Center Commentary.

Goodman, Laurie, and Jun Zhu. 2015. “Loss Severity on Residential Mortgages.” Housing Finance Policy Center Brief.

Hayre, Lakhbir. 2001. “Salomon Smith Barney guide to mortgage-backed and asset-backed securities.” Wiley and Sons 63 (2): 529–563.

Jarrow, Robert A., and Stuart M. Turnbull. 1995. “Pricing Derivatives on Financial Securities Subject to Credit Risk.” The Journal of Finance 50 (1): 53–85. Accessed April 26, 2022.

Joe, Stephen, and Frances Y. Kuo. 2008. “Constructing Sobol Sequences with Better Two-Dimensional Projections.” SIAM Journal on Scientific Computing 30 (5): 2635–2654.

Litterman, Robert B, and Thomas Iben. 1991. “Corporate bond valuation and the term structure of credit spreads.” The Journal of Portfolio Management 17 (3): 52–64. https://doi.org/10.3905/jpm.1991.409331.

Longstaff, Francis A, Sanjay Mithal, and Eric Neis. 2005. “Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market.” The journal of finance 60 (5): 2213–2253.

Longstaff, Francis A., and Arvind Rajan. 2008. “An Empirical Analysis of the Pricing of Collateralized Debt Obligations.” The Journal of Finance 63 (2): 529–563.

Netter, Andrew. 2020. “The GSE CRT market reopens post COVID-19 disruption: A new normal? Or more troubles on the horizon?” (September). https://us.milliman.com/-/media/milliman/pdfs/2020-articles/articles/9-30-20-gse-crt-market-update-v1.ashx.

O'Neill, James. 2022. "The Market View of Mortgage Credit Risk." Working Paper.

Palmer, Kevin. 2017. "What Credit Risk Transfer Tells Us About Guarantee Fees." Freddie Mac, http://www.freddiemac.com/fmac-resources/perspectives/pdf/CRT-GFee-White-Paper-Final.pdf.

Pye, Gordon. 1974. "Gauging the Default Premium." Financial Analysts Journal 30 (1): 49–52.

Richardson, Matthew P., Stijn Van Nieuwerburgh, and Lawrence J. White. 2017. "What to Do About the GSEs?" Annual Review of Financial Economics 9 (1): 21–41. https://doi.org/10.1146/annurev-financial-121415-033047.

Rowan, Tim. 1990. "Functional Stability Analysis of Numerical Algorithms." Ph.D. thesis, Department of Computer Sciences, University of Texas at Austin.

Sirignano, Justin, and Kay Giesecke. 2019. "Risk Analysis for Large Pools of Loans." Management Science 65 (1): 107–121.

Wachter, Susan M. 2018. "Credit Risk Transfer, Informed Markets, and Securitization." Economic Policy Review 24 (3).