Chapter 47
Dynamic Econometric Loss Model A Default Study of US Subprime Markets

C.H. Ted Hong

Abstract The meltdown of the US subprime mortgage market in 2007 triggered a series of global credit events. Major financial institutions have written down approximately \$120 billion of their assets to date and yet there does not seem to be an end to this credit crunch. With traditional mortgage research methods for estimating subprime losses clearly not working, revised modeling techniques and a fresh look at other macroeconomic variables are needed to help explain the crisis. During the subprime market rise/fall era, the levels of the house price index (HPI) and its annual house price appreciation (HPA) had been deemed the main blessing/curse by researchers. Unlike traditional models, our Dynamic Econometric Loss (DEL) model applies not only static loan and borrower variables, such as loan term, combined-loan-to-value ratio (CLTV), and Fair Isaac Credit Score (FICO), as well as dynamic macroeconomic variables such as HPA to project defaults and prepayments, but also includes the spectrum of delinquencies as an error correction term to add an additional 15% accuracy to our model projections. In addition to our delinquency attribute finding, we determine that cumulative HPA and the change of HPA contribute various dimensions that greatly influence defaults. Another interesting finding is a significant long-term correlation between HPI and disposable income level (DPI). Since DPI is more stable and easier to model for future projections, it suggests that HPI will eventually adjust to coincide with the DPI growth rate trend and that HPI could potentially experience as much as an additional 14% decline by the end of 2009.

47.1 Introduction

Subprime mortgages are made to borrowers with impaired or limited credit histories. The market grew rapidly when loan originators adopted a credit scoring technique like FICO to underwrite their mortgages. A subprime loan is typically characterized by a FICO score between 640 and 680 or less vs. the maximum rating of 850. In the first half of the decade, the real estate market boom and well-received securitization market for deals including subprime mortgages pushed the origination volume to a series of new highs. In addition, fierce competition among originators created various new mortgage products and a relentless easing of loan underwriting standards. Borrowers were attracted by new products such as "NO-DOC, ARM 2/28, IO" that provided a low initial teaser rate and flexible interest-only payments during the first 2 years, without documenting their income history.

As the mortgage rates began to increase during the summer of 2005 and housing activity revealed some signs of a slowdown in 2006, the subprime market started to experience some cracks as delinquencies began to rise sharply. The distress in the securitization market backed by subprime mortgages and the resulting credit crisis had a ripple effect initiating a series of additional credit crunches. All this pushed the US economy to the edge of recession and is jeopardizing global financial markets.

The rise and fall of the subprime mortgage market and its ripple effects raise a fundamental question. How can something as simple as subprime mortgages, which accounts for only 6–7% of all US mortgage loans, be so detrimental to the broader economy as well as to the global financial system?

Before formulating an answer to such a large question, we need to understand the fundamental risks of subprime mortgages. Traditional valuation methods for subprime mortgages are obviously insufficient to measure the associated risks that triggered the current market turmoil. What is the missing link between traditional default models and reality? Since a mortgage's value is highly dependent on its future cash flows, the projection of a borrower's embedded options becomes essential to simulate its cash flows. Studying consumer behavior to help project prepayments and defaults (call/put options) of a mortgage is obviously the first link to understanding the current market conditions.

This paper focuses on modeling the borrower's behavior and resultant prepayment or default decision. A Dynamic Econometric Loss (DEL) model is built to study subprime

C.H. Ted Hong (✉)
Beyondbond, Inc., USA
e-mail: ted@beyondbond.com

C.-F. Lee et al. (eds.), *Handbook of Quantitative Finance and Risk Management*,
DOI 10.1007/978-0-387-77117-5_50,

borrower behavior and project prepayment and default probabilities based on historical data from Loan Performance's subprime database (over 17 million loans) and prevailing market conditions from 2000 to 2007.

The paper is organized in the following manner. We start by constructing a general model framework in a robust functional form that is able to not only capture the impact of individual model determinants, but is also flexible enough to be changed to reflect any new macroeconomic variables. We then modeled default behavior through an individual factor fitting process. Prepayment modeling follows a similar process with consideration of the dynamic decision given prior prepayment and default history. The delinquency study builds the causality between default and delinquencies and the relationship within the spectrum of different delinquencies. We then utilized the delinquencies as a leading indicator and error correction term to enhance the predictability of the forecasted defaults by 15%. Our findings and forthcoming research are then drawn in the conclusion section (Fig. 47.1). AQ2

47.2 Model Framework

When a lender issues a mortgage loan to its borrower, the loan is essentially written with two embedded American options with an expiration co terminus with the life of the loan. The lender will then receive payments as compensation for underwriting the loan. The payments will include interest, amortized principal and voluntary/involuntary prepayments along with any applicable associated penalties. The risk for lenders is that they might not receive the contractual payments and will need to go after the associated collateral to collect the salvage value of the loan. Additionally, the foreclosure procedure could be costly and time consuming.

Unscheduled payments come in two forms. A voluntary prepayment is usually referred to simply as "prepayment" and an involuntary prepayment is known as "default" (with lags to potentially recover some portion of interest and principal proceeds). Prepayment is nothing but a call option on some or all of the loan balance plus any penalties at a strike price that a borrower has the right to exercise if the option is in-the-money. By the same token, default is a put option with the property's market value as the strike price to the borrower. Understanding the essence of both options, we need to find the determining factors that trigger a borrower to prepay/default through filtering the performance history of the loan. A list of determinant factors regarding consumer behavior theory for modeling default and prepayment will be discussed in the next two sections.

In order to construct a meaningful statistical model framework for empirical work, the availability of data and the data structure are essential. In other words, our model framework is designed to take full advantage of Loan Performance's subprime mortgage historical information and market information. The model empirically fits to the historical default and prepayment information of US subprime loan performance from 2000 to 2007 (more than 17 million loans) (Fig. 47.2). AQ3

Mathematically, our general framework constructs the default and prepayment rates as two separate functions of multiple-factors where the factors are categorized into two types – static and dynamic.[1] The static factors are initially observable when a mortgage is originated such as borrower characteristics and loan terms. Borrower characteristics include CLTV, FICO, and debt-to-income ratio (DTI). Loan terms include loan maturity, loan seasoning, original loan size, initial coupon reset period, interest only (IO) period, index margin, credit spread, lien position, documentation, occupancy, and loan purpose. The impact to the performance of a loan from the static factors provides the initial causality impact, yet their influence may diminish or decay as the information is no longer up to date.

Dynamic factors include several macroeconomic variables such as HPA, prevailing mortgage interest rates, consumer confidence, gross disposable income, employment rate, and unemployment rate. These dynamic factors supply up-to-date market information and thus play an important role in dynamically capturing market impact. The accuracy of capturing causality impact due to the static factors and the predictability of the dynamic factors presented a constant challenge during the formulation of this model.

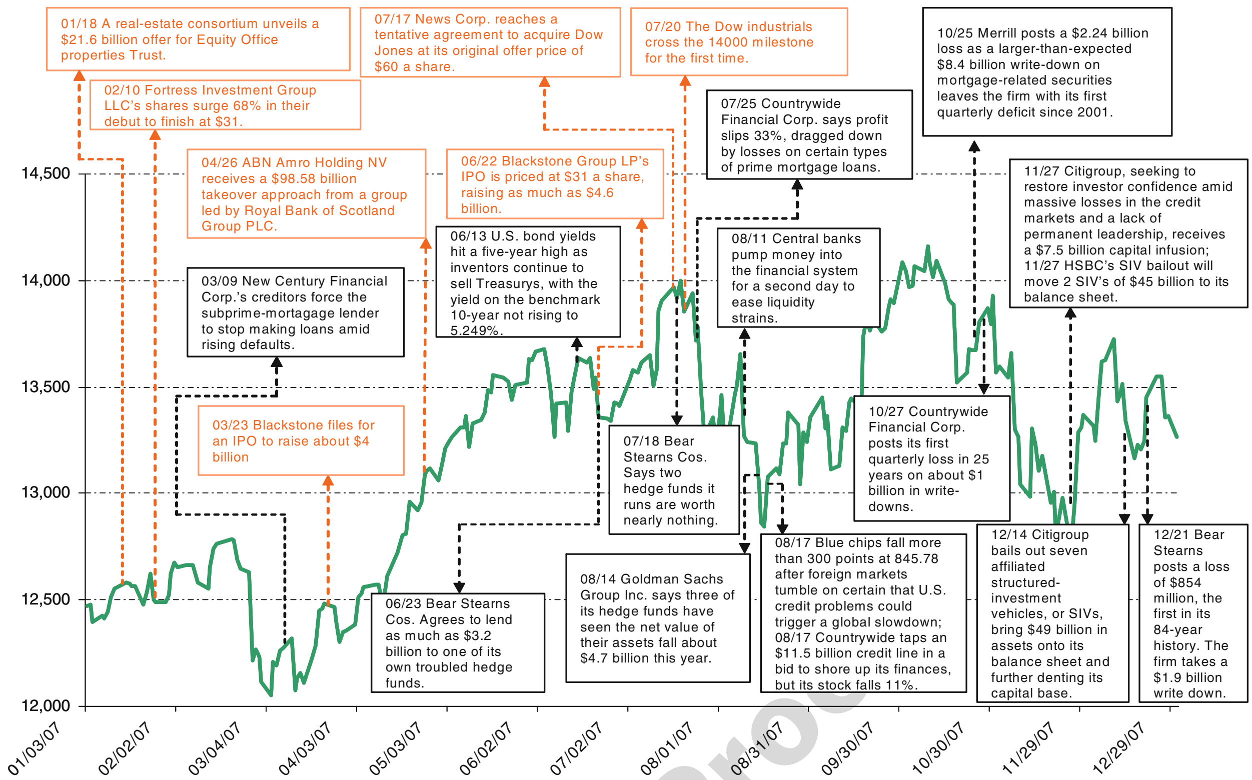
For each individual factor, a non-linear function is formulated according to its own characteristics. For example, a "CLTV" factor for modeling default is formulated as the function of default rate over CLTV ratio. However, a DOC factor is formulated as the function of multiplier over discrete variables of "FULL" vs. "LIMITED" with percentages of respective groups.

A general linear function of combined multifactor functions is then constructed as a basic model framework to fit the empirical data and to project forecasts for prepayments and defaults.[2] In the following sections, we will discuss each factor in detail.

[1] There is no industry standard measure for default rate, thus a different definition on default rate will give a very different number. As there is no set standard, we define our default rate based on the analysis in this paper, "Loss Severity Measurement and Analysis," The MarketPulse, LoanPerformance, 2006, Issue 1, 2–19. Please refer to Appendix I for definition of default used throughout this paper.

[2] See Appendix II for the details of model specification.

BookID 149174_ChapID 050_Proof# 1 - 17/12/09



What happened in 2008

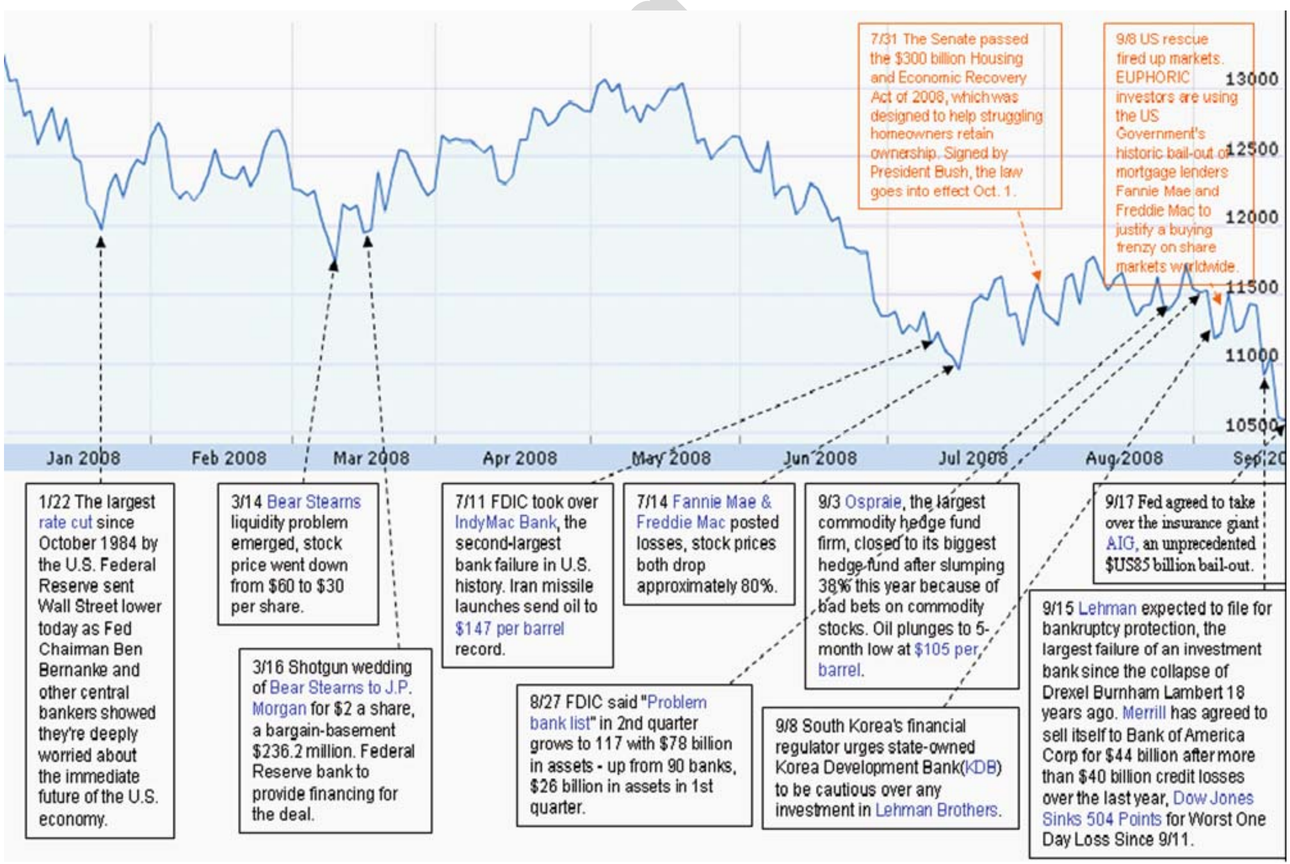


Fig. 47.1 What happened in 2007; What happened in 2008

Type / Orig. Year	ARM OTHER	ARM2/28	ARM3/27	ARM5/25	FIXED	Grand Total
2000	11,452	187,232	68,430	4,059	390,671	661,844
2001	11,389	261,316	67,018	10,449	477,718	827,890
2002	33,776	434,732	100,939	25,827	605,233	1,200,507
2003	51,548	697,073	164,228	71,839	958,170	1,942,858
2004	221,818	1,239,522	413,366	213,572	1,172,413	3,260,691
2005	496,697	1,577,003	393,020	301,829	1,619,257	4,387,806
2006	490,975	1,137,345	234,344	349,460	1,754,382	3,966,506
2007	99,946	161,480	36,795	160,549	404,278	863,048
Grand Total	1,417,601	5,695,703	1,478,140	1,137,584	7,382,122	17,111,150

Fig. 47.2 Number of securitized Alt-A and subprime mortgage origination

47.3 Default Modeling

Default Modeling Factor Components

Seasoning	Occupancy
Combined Loan-to-Value (CLTV)	Owner
Credit Score (FICO)	Second home
Debt-to-Income Ratio (DTI)	Investor
Payment Shock (IO)	Property Type
Relative Coupon Spread	Single-Family
Loan Size	Multi-Family
Lien	Condo
First	Loan Documentation
Second and Others	Full
Loan Purpose	Limited
Purchase	House Price Appreciation (HPA)
Refinance	State Level
Cashout	CBSA Level

47.3.1 Seasoning

Loan information regarding borrower's affordability is usually determined at origination. As a loan seasons, its original information decays and its default probability starts to surge. A seasoning baseline curve with annualized Constant Default Rate (CDR) against its seasoning age would post a positive slope curve for the first 3 years.

Figure 47.3 shows actual CDR curves and their fitted result of different vintages of ARM 2/28 mortgage pools. They roughly follow a similar shape to the Standard Default Assumption (SDA) curve.[3] However, as shown in Fig. 47.4, the ramp-up curve can be very different for different vintages.

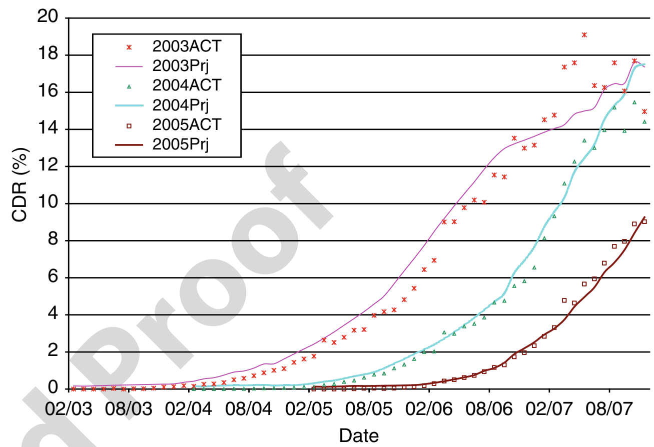
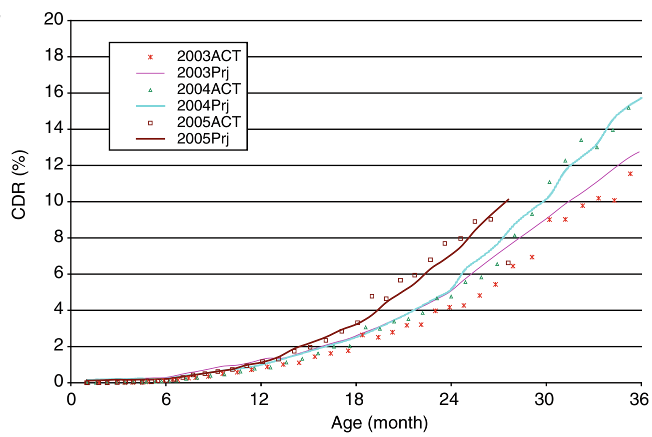


Fig. 47.3 Seasoning: CDRs by date and vintages of ARM 2/28



Source: Beyondbond Inc, LoanPerformance

Fig. 47.4 Seasoning: CDRs by age and vintages of ARM 2/28 (Source: Beyondbond Inc, LoanPerformance)

47.3.1.1 Why Is the 2005 Seasoning Pattern Faster Than Prior Vintages?

Since the seasoning baseline curve is not independent of dynamic factors, a dynamic factor such as HPA could tune vintage seasoning curves up and down. In Fig. 47.4, the 2005 seasoning pattern is significantly steeper than its prior

[3] SDA is based on Federal Housing Administration (FHA)'s historical default rate and was developed by Bond Market Association (BMA), now known as Securities Industry and Financial Markets Association (SIFMA).

vintages. Looser underwriting standards and deteriorating credit fundamentals can be an important reason. Negative HPA obviously starts to adversely impact all vintages after 2005.

47.3.2 Payment Shock – Interest Only (IO)

The boom in the subprime market introduced new features to the traditional mortgage market. An ARM 2/28 loan with a 2-year interest-only feature has a low fixed initial mortgage rate and also pays no principal for the first 2 years prior to the coupon reset.[4]

When the IO period ends, the borrower typically faces a much higher payment based on its amortized principal plus the fully indexed interest. This sudden rise in payments could produce a "Payment Shock" and test the affordability to borrowers. Without the ability to refinance, borrowers who are either under a negative equity situation or not able to afford the new rising payment will have a higher propensity to default. Consequently, we see a rapid surge of default rates after the IO period.

The ending of the IO period triggers payment shock and will manifest itself with a spike in delinquency.[5] Delinquent loans eventually work themselves into the defaulted category within a few months after the IO period ends. Figure 47.5 shows the different patterns and the default lagging between IO and Non-IO ARM 2/28 pools.

47.3.3 Combined Loan-to-Value (CLTV)

LTV measures the ratio of mortgage indebtedness to the property's value. When multiple loans have liens added to the indebtedness of the property, the resulting ratio of CLTV becomes a more meaningful measure of the borrower's true equity position.

However, the property value might not be available if a "market" property transaction does not exist. A refinanced mortgage will refer to an "appraisal value" as its property value. Note that "appraisal value" could be manipulated during ferocious competition among lenders in a housing boom market and would undermine the accuracy of CLTV.

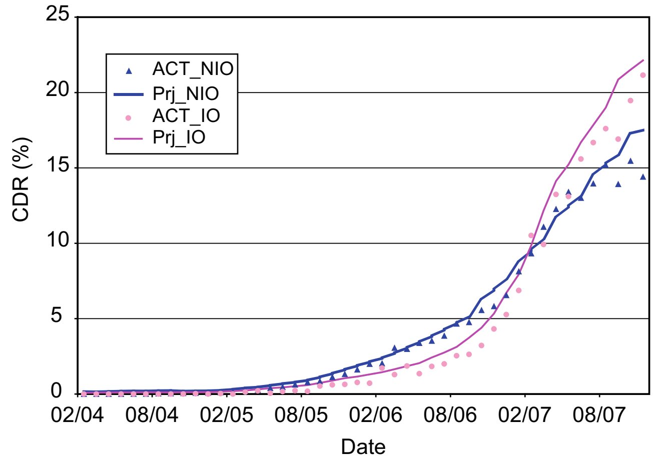


Fig. 47.5 IO payment shock: CDRs by date of ARM 2/28

As we know, default is essentially a put option embedded in the mortgage for a borrower. In a risk neutral world, a borrower should exercise the put if the option is in-the-money. In other words, a rational borrower should default if the CLTV is greater than one or the borrower has negative equity.

At higher CLTVs, it becomes easier to reach a negative equity level as the loan seasons and its default probability increases. Figure 47.6 provides the actual stratification result of CDR over various CLTV ranges. Obviously, CDR and CLTV are positively correlated. In addition, lower CDR values are observed for higher subprime tiered FICO ranges. This shows that the FICO tier granularity is another important factor in modeling.

However, since CLTV is obtained at the loan's origination date, it does not dynamically reflect housing market momentum. We introduce a dynamic CLTV that includes housing price appreciation from loan origination in order to estimate more precisely the actual CLTV. This dynamic CLTV allows us to better capture the relationship between CLTV and default. Figure 47.7 clearly illustrates that different CLTV groups show a different layer of risk level.

47.3.4 FICO

FICO score is an indicator of a borrower's credit history. Borrowers with high FICO scores maintain a good track record of paying their debts on time with a sufficiently long credit history.[6]

[4] The reset is periodical, and the interest rate is set as Index + Margin.

[5] The delinquency rate is measured by OTS (Office of Thrift Supervision) or MBA (Mortgage Bankers Association) convention. The difference between these two measures is how they count missed payments. MBA delinquency rate counts the missed payment at the end of the missing payment month while OTS delinquency rate counts the missed payment at the beginning of the following month after missing payment. This difference will pose a 1–30 days delay of record. OTS delinquency rate is the prevailing delinquency measure in subprime market.

[6] According to Fair Isaac Corporation's (The Corporation issued FICO score measurement model) disclosure to consumers, 35% of this score is made up of punctuality of payment in the past (only includes pay-

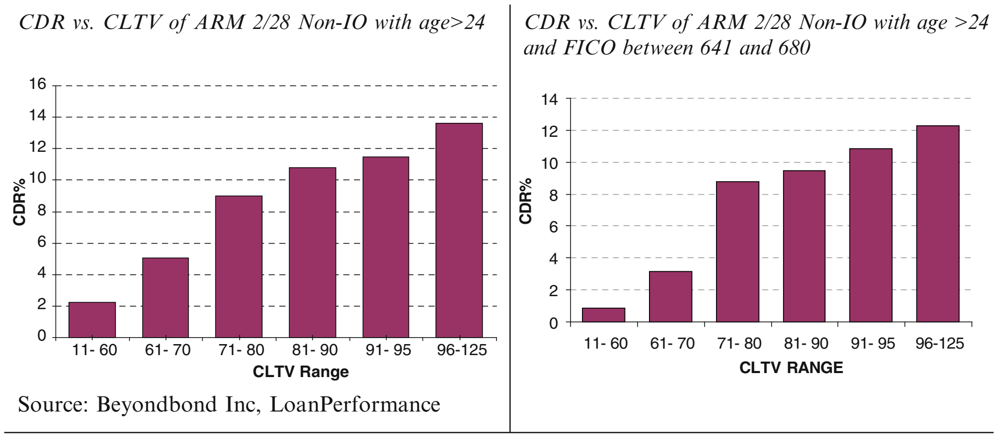
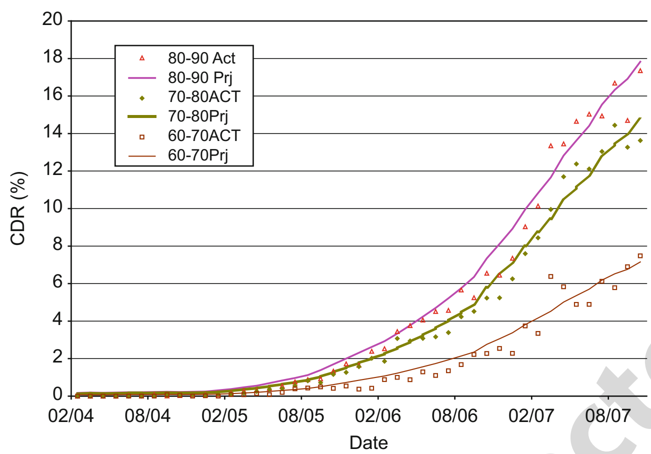


Fig. 47.6 Stratified seasoned CDR over CLTV ranges (Source: Beyondbond Inc, LoanPerformance)



Source: Beyondbond Inc, LoanPerformance

Fig. 47.7 CDRs by date and CLTVs of ARM 2/28 (Source: Beyondbond Inc, LoanPerformance)

In recent years, people came to believe that FICO was no longer an accurate indicator due to the boom in hybrid ARM loans and fraudulent reporting to the credit bureaus. Since refinancing was much easier to obtain, issuers were giving out tender offers to borrowers in order to survive the severe competition among lenders.

CLTV and FICO scores are two common indicators that the industry uses to predict default behavior.[7] We examine the combined CLTV and FICO effects on CDR as shown in Fig. 47.8. The figure presents a 3-D surface of stratified CDR rates over CLTV and FICO ranges from two different angles for seasoned ARM 2/28 pools. The relationship between CLTV and CDR is positively correlated across various FICO ranges. On the other hand, the relationship between FICO and CDR is somewhat negatively correlated across various CLTV ranges. However, the case is not as significant. FICO's impact is obviously not as important as we originally expected.

In our analysis, CLTV $=$ 75 and FICO $=$ 640 serves as a base curve, and then we adjust the CDR according to movements of other default factors.

Figure 47.9 gives an example of fitting results based on ARM 2/28 2004 vintage pools. The difference between 600–640 and 680–700 FICO ranges makes only a small difference of 1% in CDR for a seasoned pool.

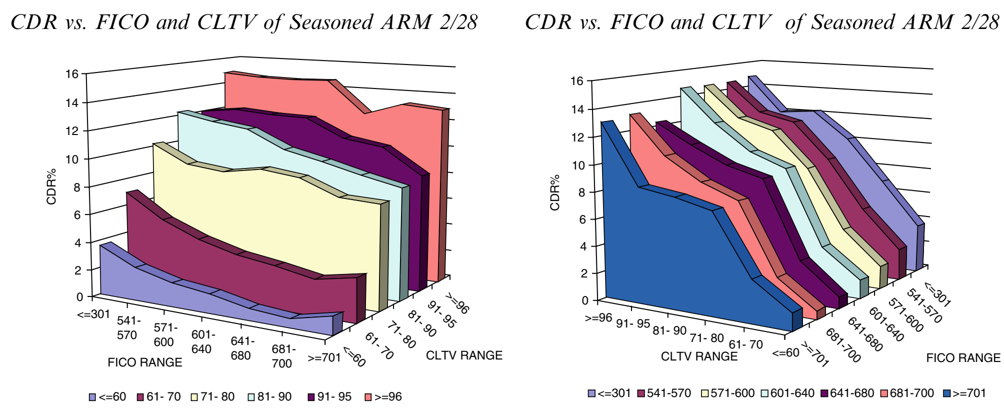
47.3.5 *Debt-to-Income Ratio (DTI) and Loan Documentation (DOC)*

The DTI in this paper is defined as the back-end DTI, which means the debt portion for calculating the DTI ratio includes not only PITI (Principal + Interest + Tax + Insurance) but also other monthly debts such as credit card payments, auto loan payments and other personal obligations.[8] The DTI ratio shows the affordability of a loan to a borrower and provides us with a clearer picture of a borrower with an exceptionally

ments later than 30 days past due), 30% is made up of the amount of debt, expressed as the ratio of current revolving debt (credit card balances, etc.) to total available revolving credit (credit limits), and 15% is made up of length of credit history. Severe delinquency (30 plus) and credit history length make up 50% of the FICO score. This score reflects people's willingness to repay. It's essentially the probability distribution for people's default activity on other debts such as credit card and/or utility bills, etc. Statistically speaking, people with higher FICO scores will have lower probability to default.

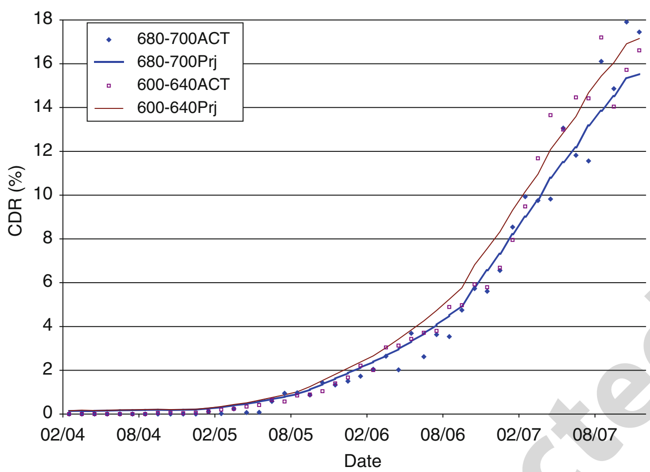
[7] Debt-to-Income ratio is also an important borrower characteristic, but in recent years, more Limited-Doc or/and No-Doc loans are issued. For these loans, many of them do not have DTI ratio report, so we consider DTI separately for different DOC type.

[8] There are two major measures of DTI in the industry: Front-End DTI ratio = PITI/Gross Monthly Income, and Back-End-DTI ratio = PITI + Monthly Debt/Gross Monthly Income. PITI = Principle + Interest + Tax + Insurance.



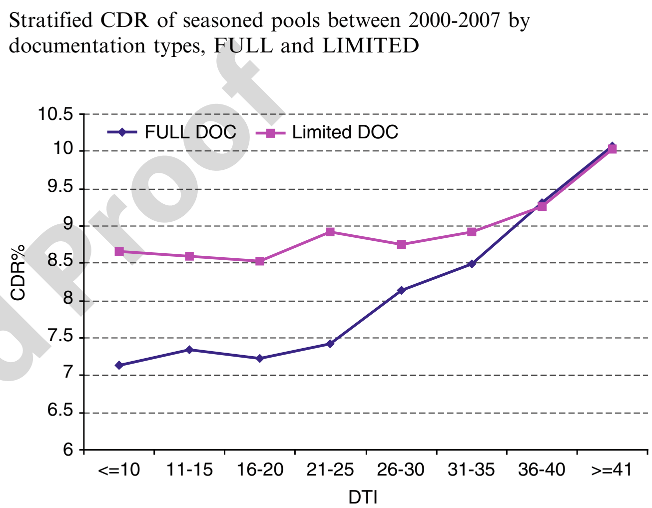
Source: LoanPerformance

Fig. 47.8 Stratified CDR by CLTV and FICO of ARM 2/28 (Source: LoanPerformance)



Source: Beyondbond

Fig. 47.9 FICO: CDRs by date and FICOs of ARM 2/28 (Source: Beyondbond)



Source: Beyondbond Inc, LoanPerformance

Fig. 47.10 Stratified CDR by DTI ranges (Source: Beyondbond Inc, LoanPerformance)

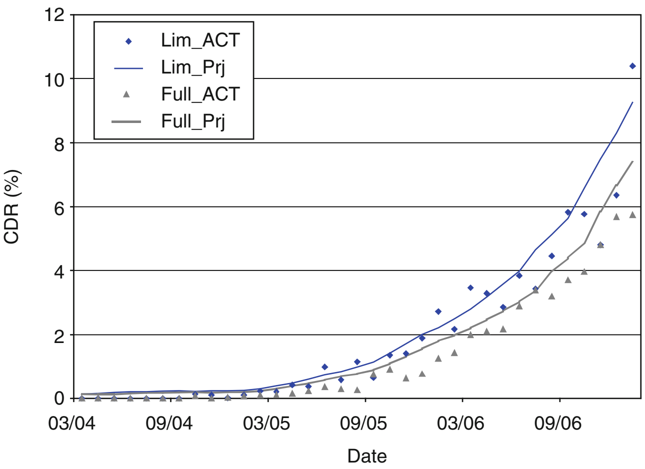
high DTI. For different regions of the country, the DTI ratio could imply a different financial condition of the borrower because of different living standards and expenses between those of rural areas and large cities.

DTI is captured and reported as part of the loan documentation process. Loan documentation, also referred to as DOC, consists of three major groups: "FULL DOC," "LOW DOC," and "NO DOC." Lenders usually require a borrower to provide sufficient "FULL" documentation to prove their income and assets when taking out loans. People who are self-employed and/or wealthy and/or have lumpy income stream are considered as borrowers with "LIMITED" (LOW or NO) documentation. In recent years, fierce competition pushed lenders to relax their underwriting standards and originated more LIMITED DOC loans with questionable incomes. This uncertainty regarding income poses uncertainties in determining the real DTI.

The stratification report shows two very different patterns of default between FULL and LIMITED documentation categories when analyzing the DTI effect. For FULL DOC loans, default probability vs. DTI is very much positively correlated, CDR increases as the DTI increases. Since FULL DOC loans are loans that have documented income and assets, it shows the default DTI relationship most clearly in Fig. 47.10. LIMITED DOC has a weaker relationship compared to FULL DOC. Figure 47.11 shows the two different time series patterns of CDR curves and their fitted values between FULL and LIMITED DOCs.

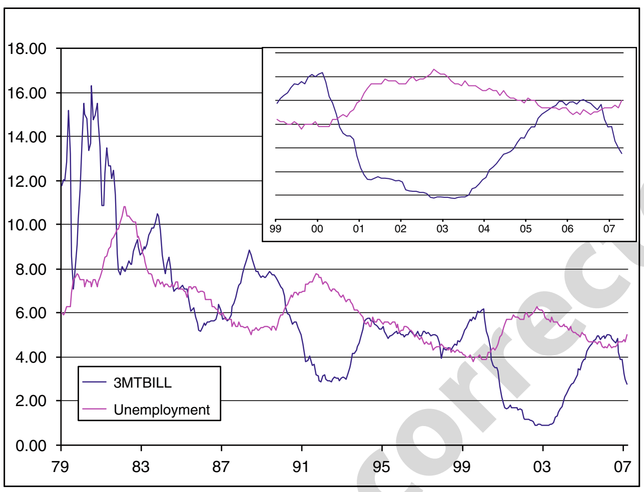
Since income is one of the main elements in determining the DTI ratio, the macroeconomic variable, unemployment rate, becomes an important determinant that affects an individual's income level. We found an interesting result when we plotted the unemployment rate against 3-month US

Actual versus Fitted CDR curve over time by documentation types, FULL and LIMITED for 2004 vintages



Source: Beyondbond Inc, LoanPerformance

Fig. 47.11 DOC: actual vs. fitted for 2004 (Source: Beyondbond Inc, LoanPerformance)



Source: US Bureau of Census

Fig. 47.12 Unemployment rate 1979–2007 (Source: US Bureau of Census)

Source: US Bureau of Census

Fig. 47.13 Unemployment vs. T-bill 3 months (Source: US Bureau of Census)

Treasury Bills. They have been very negatively correlated for the last 7 years. Whether it was a coincidence or not, it suggests that the monetary policy has been mainly driven by the AQ4 unemployment numbers (Figs. 47.12 and 47.13).

47.3.6 Loan Size

Is bigger better? The conventional argument is that larger loan size implies a better financial condition and lower likelihood of default. According to the stratification results based on original loan size in Fig. 47.14, CDR forms a smile curve across original loan balance. Loans with amounts larger than \$350,000 tend to be a bit riskier although the increment is marginal. Loans with an amount less than \$100,000 also seem riskier. Larger loans do not seem to indicate that they are better credits. The original loan size is usually harder to interpret as it can be affected by other factors such as lien, property type, and geographical area. For example, a \$300,000 loan in a rural area may indicate a borrower with growing financial strength; while the same amount in a prosperous large city may indicate a borrower with weak purchasing power. Without putting size into the context of property type and geographic location, the factor could be misleading. This may explain why we do not see a clear shape forming in Fig. 47.14. Figure 47.15 shows the three different time series patterns of CDR curves and their fitted values based on their loan size ranges. Since the size is mixed for all the property types, the pattern and fitted results for each category are distorted and the fit is not as good as other factors.

47.3.7 Lien

We know that a second mortgage/lien has a lower priority to the collateral asset than a first lien mortgage/lien in the event of a default. Thus, the second lien is riskier than the first lien. Second lien borrowers usually maintain higher credit scores, typically with a FICO greater than 640. We often see a very mixed effect if this layered risk is not considered. In Fig. 47.16, second lien loans are significantly riskier than first lien loans when measured against comparable FICO ranges for both liens. Figure 47.17 shows two different time series patterns of CDR curves and their fitted values based on their liens.

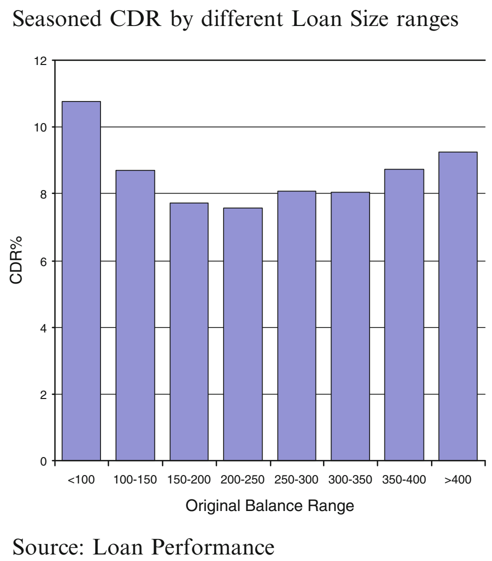


Fig. 47.14 Loan size stratification (Source: LoanPerformance)

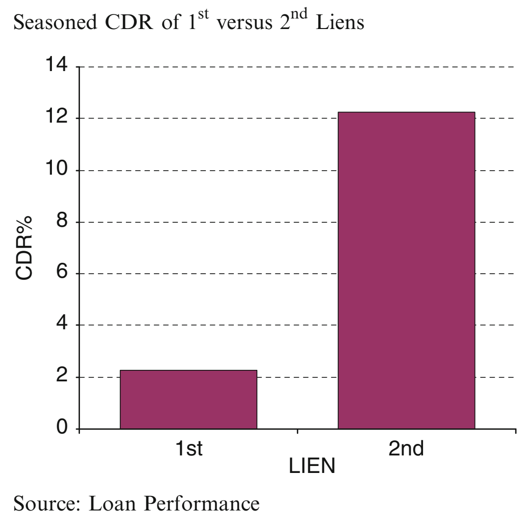


Fig. 47.16 Lien stratification (Source: LoanPerformance)

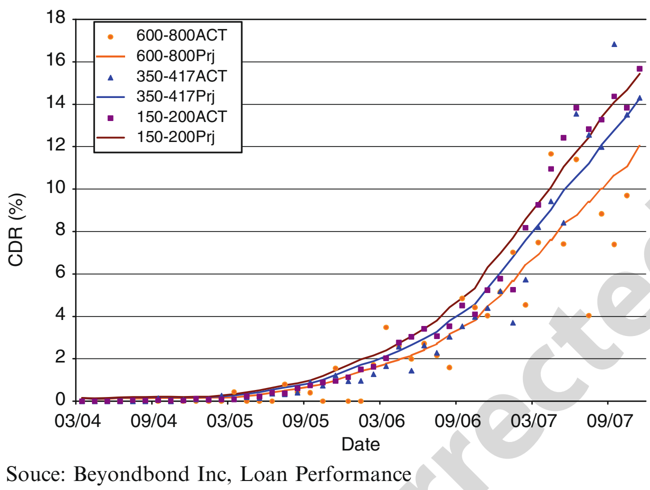


Fig. 47.15 Size: actual vs. fitted CDR for 2004 (Source: Beyondbond Inc, LoanPerformance)

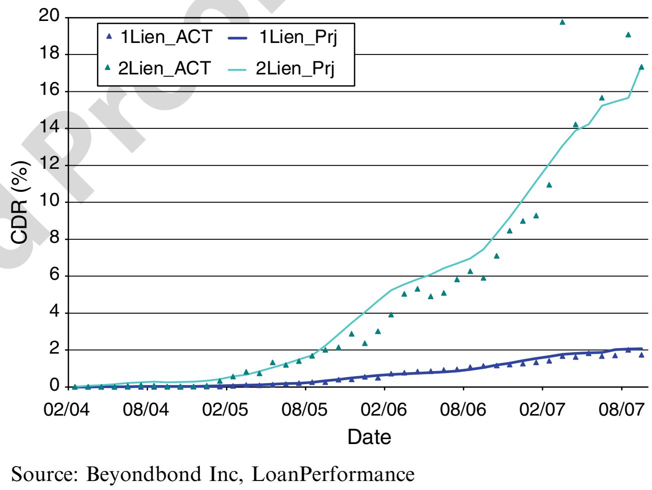


Fig. 47.17 Lien: actual vs. fitted CDR for 2004 (Source: Beyondbond Inc, LoanPerformance)

47.3.8 Occupancy

Occupancy consists of three groups: “OWNER,” “INVESTOR,” and “SECOND HOME.” The “OWNER” group views the property as their primary home, rather than as an alternative form of housing or an investment. This group will face emotional and financial distress if the property is in foreclosure or REO. Thus, this group has a lower propensity to default compared with others if all other factors remain the same. On the other hand, “INVESTOR” and “SECOND HOME” groups would be more risk neutral and are more willing to exercise their options rationally. In other words, they should have a higher default risk.

Figure 47.18 reports an occupancy stratification regarding the default risk profile. The result evidently supports the risk neutral idea with respect to the “INVESTOR” group and “INVESTOR” does show the highest default risk among all three groups. The “OWNER” group, however, is not the lowest default risk group. Instead, the “SECOND HOME” group is the lowest one. The observation is interesting, but not intuitive. It indicates that when a borrower faces financial stress, a “SECOND HOME” will be sold first even at a loss to support his/her primary home. Thus, the default risk of “SECOND HOME” is actually reduced by incorporating a borrower’s primary home situation and cannot be simply triggered by the risk neutral idea. Figure 47.19 shows the two different time series patterns of CDR curves and their fitted values between “OWNER” and “INVESTOR.”

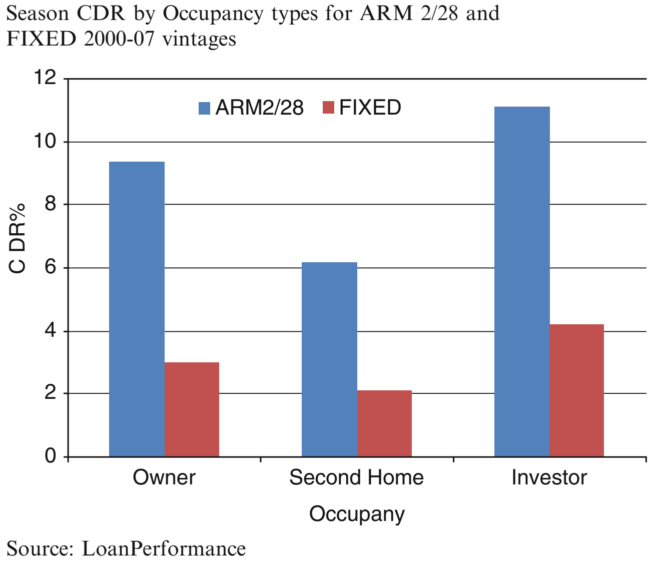


Fig. 47.18 Occupancy stratification (Source: LoanPerformance)

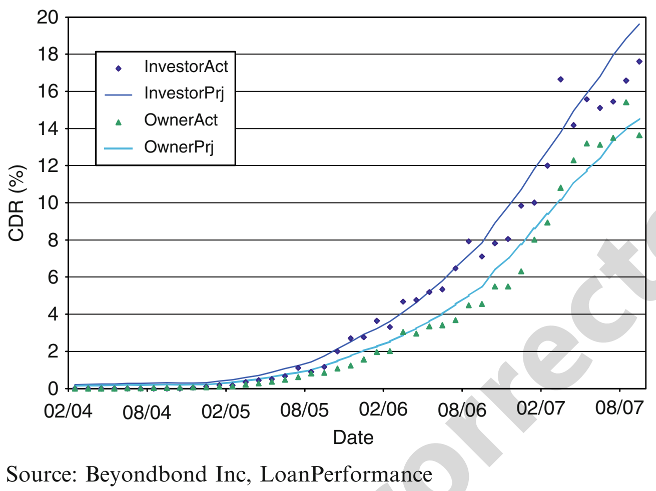


Fig. 47.19 Occupancy: actual vs. fitted CDR for 2004 (Source: Beyondbond Inc, LoanPerformance)

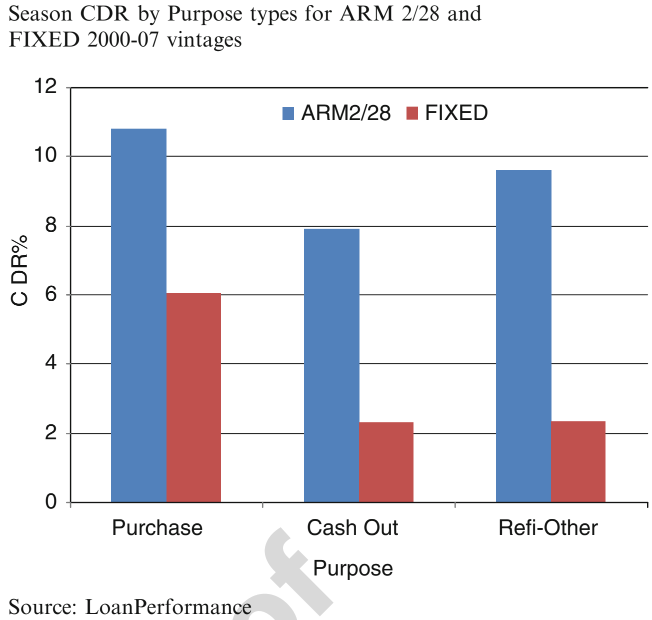


Fig. 47.20 Purpose stratification (Source: LoanPerformance)

47.3.9 Purpose

Loan Purpose classifies three key reasons for receiving a loan as "PURCHASE," "CASHOUT," and "REFI."[9] "PURCHASE" means the borrower is receiving his/her first loan on the property. "REFI" uses the loan for refinancing the outstanding balance without any additional funds drawn from the equity in the property. "CASHOUT" refers to a refinance loan with extra cash inflow to the borrower due to the difference between the new increased loan amount and the existing loan balance (Fig. 47.20).

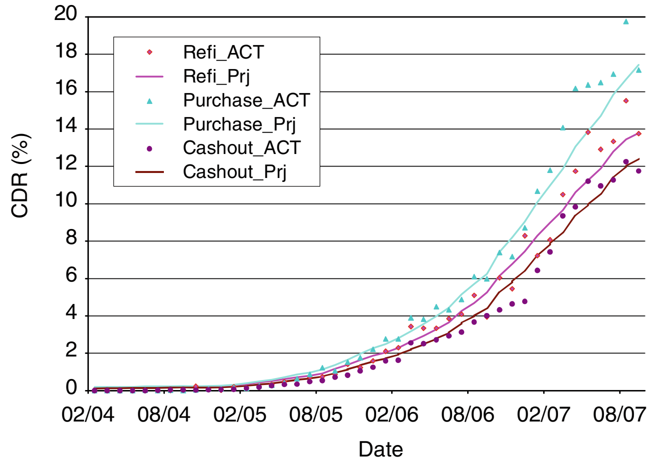
"CASHOUT" and "REFI" usually reflects an intention to rollover the IO period or benefit from a lower mortgage rate. They can only be afforded by borrowers in good financial condition. "REFI" is a group of borrowers with a higher FICO, LTV as compared to the other two categories. So we expect the "REFI" loans to have a lower default rate than "PURCHASE" loans. The argument seems correct for fixed rate mortgages. "REFI" borrowers have much lower default probability than "PURCHASE."

Beginning in 2007, the credit crunch hit the market and most of the lenders tightened their credit standards. Hybrid ARM loans, such as an ARM 2/28, faced new resets and borrowers who no longer qualified for refinancing were in danger of defaulting. If these people can no longer afford the payment after IO and/or reset, they will eventually enter default. ARM 2/28 loans show a significant increase in defaults for "REFI" purpose as compared with FIXED rate loans. Figure 47.21 shows the three different time series patterns of CDR curves and their fitted values among various purpose types.

47.3.10 Dynamic Factors: Macroeconomic Variables

As we mentioned in the model framework, macroeconomic variables such as HPI, interest rate term structure, unemployment rate, and others that supply up-to-date market information can dynamically capture market impact.

[9] For simplicity sake, we categorize refinance, second mortgage, and other miscellaneous types as "REFI."



Source: Beyondbond Inc, LoanPerformance

Fig. 47.21 Purpose: actual vs. fitted CDR for 2004 (Source: Beyondbond Inc, LoanPerformance)

In theory, an economy maintains its long-term equilibrium as "Norm" in the long run and a handful of macroeconomic variables are usually used to describe the situation of the economy. While the economy is in its "Norm" growing stage, these macroeconomic variables usually move or grow very steadily and the risk/return profile for an investment instrument can be different depending on its unique investment characteristics. Because of that, a diversified investment portfolio can be simply constructed based on relationships in the correlation matrix. Thus, macroeconomic variables are usually ignored during the "Norm" period. However, when an economy is under stress and approaches a "bust" stage, many seemingly uncorrelated investments sync together. The same macro variables become the main driving forces that crucially and negatively impact the investment results. The current credit crunch is creating mark to market distress for investments across not only various market sectors but also credit ratings, which clearly describes our view regarding these macroeconomic variables.

Since the severe impact from these variables typically occurs in economic downturns, cross correlation could provide a preliminary result in understanding the causality and the magnitude of their relationship. The dynamic interaction between these variables and consumer behavior would then provide a better sense of prediction and therefore either prevent the next downturn or efficiently spot an investment opportunity based on the next market recovery.

47.3.11 House Price Appreciation (HPA)

AQ6 The house price index (HPI) has been the most quoted macroeconomic variable that measures or determines high delinquency and default rates since the beginning of subprime crisis.[10] Thus, house price appreciation (HPA), which measures the housing appreciation rate, year-over-year, has become the most important indicator within the US housing market. By comparing the 30-day delinquency across vintages, we see that delinquency rates increase after the 2005 vintage (Fig. 47.22).

When we look at seasoning patterns across 2000–2005 vintages, we find that the 2005 seasoning pattern started to surge after 18 months of age or the third quarter of 2006. Coincidentally, HPA started to decline in the second quarter of 2006. Although a similar HPA pattern appeared at the third quarter of 2003, the main difference was that the former one was the up-trend of HPA, while the latter was on a down-trend. Defaults in 2003 were obviously lower than in 2006 with comparable loan features and seasoning/age. In order to capture this subtle trend difference, we studied HPI and its various dimensions in addition to HPA levels (Fig. 47.23). AQ7

47.3.11.1 Multi-Dimension HPI Impacts

To systematically identify the impact of HPA, we measure HPA in three aspects regarding each loan:

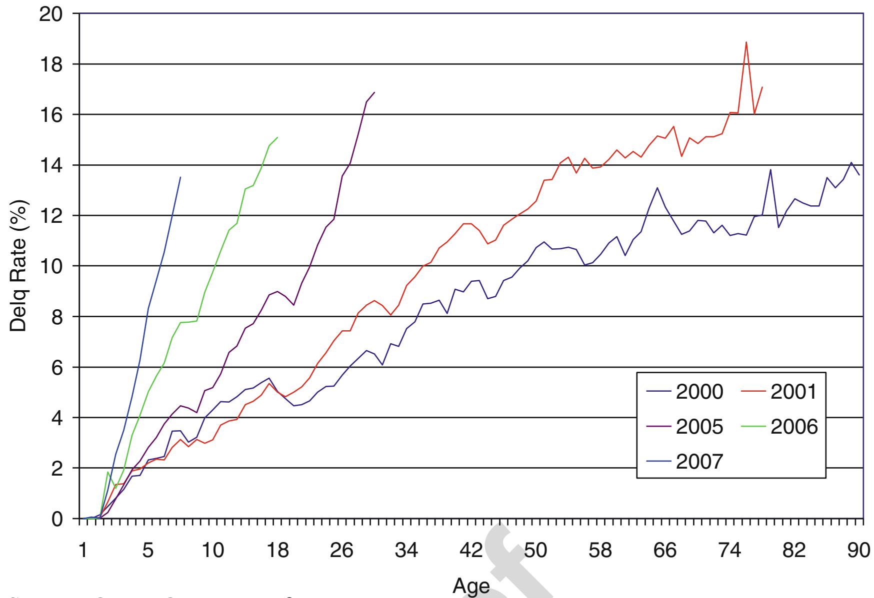
- Cumulative HPI, an accumulative HPA since origination, is calculated based on HPI levels to capture equity gain for borrowers.
- HPA, the change rate of HPI, captures the pulse of the housing market.
- HPA2D, the change of HPA, is used to capture the trend/expectations of the housing market.

HPA factors form a multi-dimensional impact to reflect a loan's up-to-date capital structure, current housing market conditions, and future housing market prospects. We embedded the "Cumulative HPI" into CLTV to build a dynamic CLTV to reflect the dynamic equity value to the property. In a risk neutral analysis, an option model can be easily applied to project the default probability. HPA is already a leading market indicator in explaining defaults. HPA2D basically serves as the second derivative of HPI; it allows us to capture the general expectation on home price movements and market sentiment.

The negative impact due to HPA2D in the third quarter of 2006 is apparently different from the third quarter of

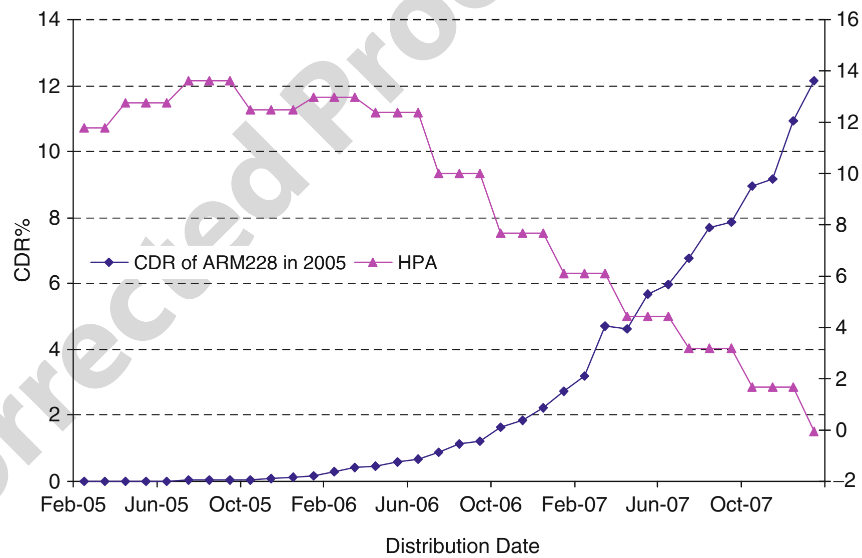
[10] The Housing Price Index (HPI) used in this paper is published by Office of Federal Housing Enterprise Oversight (OFHEO) as a measure of the movement of single-family house prices. According to OFHEO, The HPI is "a weighted, repeat-sales index, meaning that it measures average price changes in repeat sales or refinancing on the same properties." See website of OFHEO www.ofheo.gov for details.

Fig. 47.22 30-Day delinquency of ARM2/28 by vintages (Source: OFHEO, LoanPerformance)



Source: OFHEO, LoanPerformance

Fig. 47.23 HPA vs. CDR of ARM2/28 2005 vintage (Source: OFHEO, LoanPerformance)



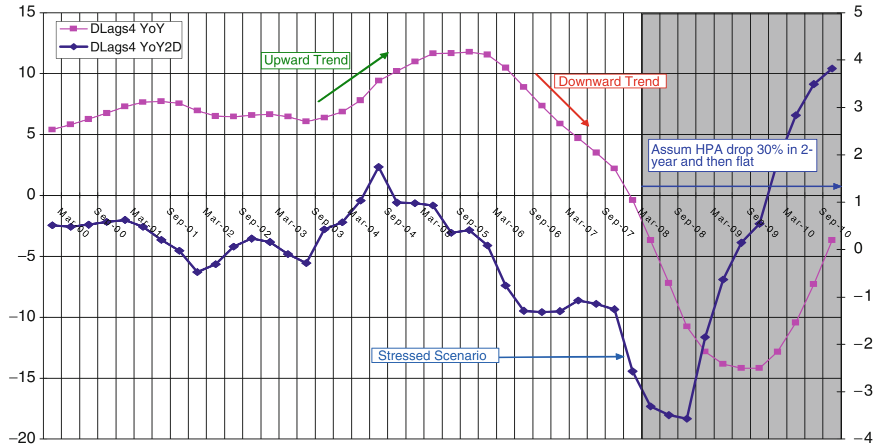
Source: OFHEO, LoanPerformance

2003 even though the HPA numbers are at a similar level.[11] HPA2D undoubtedly offers another dimension that reflects consumer expectations about the general housing market. When HPA2D is negative, the probability of borrowers holding negative equity increases.

The remaining challenge lies in the deterioration of the housing market, which is producing unseen record-low HPI levels. While the HPA continues decreasing, HPA2D plunges even faster. Our multi-dimensional HPA empirical fitting merely relies on a very limited range of in-sample HPA data. To extrapolate HPA and HPA2D requires numerous possible market simulations to induce a better intuitive sense of the numbers. The shaded area in Fig. 47.24 shows a simulated extreme downturn in the housing market that assumes a 30% drop in HPI levels based on the fourth quarter of 2007 and then a leveling-off. Based on the simulation results, HPA2D starts to pick up at least one-quarter earlier than HPA and 1-year earlier than HPI. While a 2-year HPI downturn is assumed, the consumers' positive housing market expectation reflected in HPA2D effectively reduces their incentive to walk away from their negative equity loans. This case example clearly shows how the forecasted HPA and HPA2D numbers could provide a better intuitive market sense to the model.

[11] We have smoothed HPA and HPA2D series to create better trend lines. A linear weighted distributed lags of last four quarters are adopted for smoothing the series.

Fig. 47.24 HPA vs. HPA2D, actual and extreme simulation (Source: OFHEO, LoanPerformance)



Source: OFHEO, Loan Performance

The relationship between HPI and consumer behavior that forms the HPI impact to default and prepayment are then modeled. We illustrate the multi-dimensional HPI impact through an example shown below:

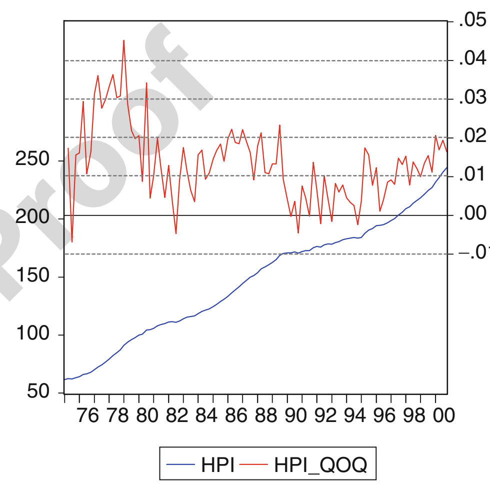
1. HPCUM ↓(below 5%) ⇒ CLTV↑ ⇒ MDR↑, SMM ↓
2. HPA ↓(below 2%) ⇒ MDR↑, SMM ↓
3. HPA2D ↓(below −5%) ⇒ MDR↑, SMM ↓

47.3.11.2 HPI and DPI

When the economy is experiencing a potentially serious downturn, generating HPI predictions going out 3 years is a much better approach than random simulations. Since HPI has increased so rapidly since 2000, the current decline could be merely an adjustment to the previously overheated market. The magnitude and ramp-up period of the adjustment nevertheless determines consumers' behavior of exercising their mortgage embedded options. Finding a long-term growth pattern of HPI thus becomes very AQ8 vital for predicting and simulating future HPI numbers (Fig. 47.25).

Based on Fig. 47.26, HPI draws a constant relationship with disposable personal income (DPI) in the long run. Since DPI is a more stable process, a long-term HPI prediction based on the observed relationship between DPI and HPI provides a better downturn average number. Based on our long-term HPI prediction, HPI could potentially drop as much as 14% by the end of 2009.[12]

[12] A 5% decrease by the end of 2009 on average plus another 9% based on two standard errors of regression of HPI on DPI result.



Source: US Bureau of Census

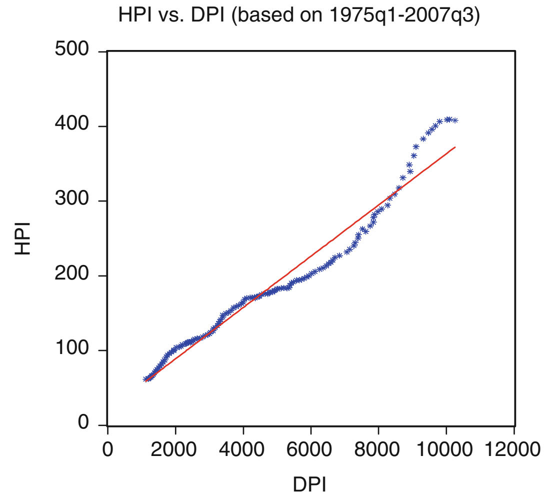
Fig. 47.25 HPI & HPA QoQ 1975–2007 (Source: US Bureau of Census)

47.3.11.3 Geographical Location and Local HPI

In housing markets, geographic location (location, location, location or L^3) is undoubtedly the most important price determinant as it is globally unique. While we are pointing out all HPI impacts in general, HPI at the national level does not reflect the actual local situation and thus distorts the default impact ignoring the granularity of detailed local housing market information. The consequence of ignoring this kind of granularity can be very severe when a geographically diversified mortgage pool's CLTV has a fat-tailed distribution in its high CLTV end. Since local HPI can vary from national HPI, loans with negative equity have a higher level of relevance than the use of national HPI.

Fortunately, we are able to differentiate HPI impact by drilling down to the state as well as to the CBSA level.

Fig. 47.26 HPI vs. DPI (Source: US Bureau of Census)



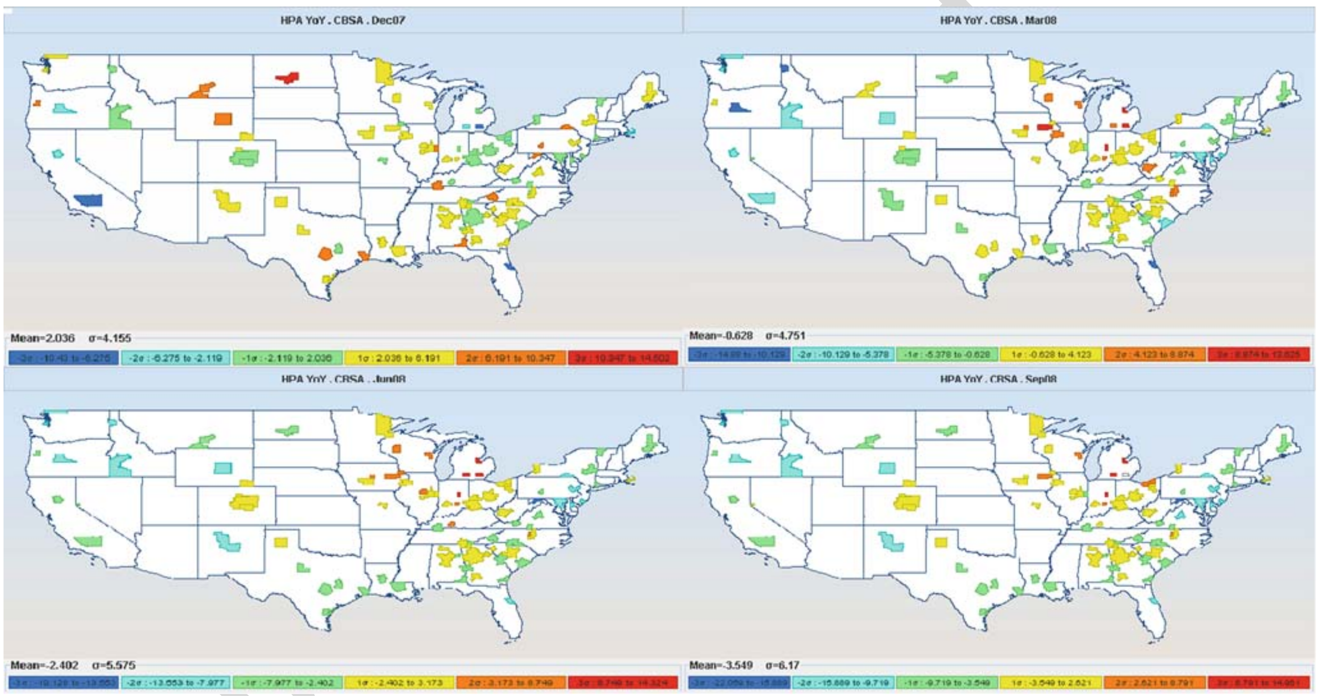


Fig. 47.27 Geographic components of HPA by CBSA

Figure 47.27 shows the actual levels of HPA on December 2007 and our projection of HPA for June 2008 detailed by CBSA. We started with a national level HPI model to obtain the long-term relationship between HPI and DPI. We then build a dynamic correlation matrix between national and state as well as national and CBSA levels that dynamically estimates parameters and generates forecasts on the fly. The CBSA level HPI is especially important for calculating dynamic CLTV. Since the cumulative HPI (HPCUM) is calculated as the cumulative HPA since origination, it captures the wealth effect for generating dynamic CLTVs. This more detailed information helps to predict if a mortgage has crossed into the negative equity zone.

47.4 Prepayment Modeling

Prepayment Modeling Factor Components
Housing Turnover and Age
Refinancing

Teaser Effect
Interest Only (IO) Effect
Burnout Effect
Seasonality
Loan-to-Value Effect
FICO Credit Effect
Prepayment Penalty
House Price Wealth Effect

47.4.1 Housing Turnover and Seasoning

Housing turnover rate is the ratio of total existing single-family house sales over the existing housing stock.[13] With the exception of cases in the early 1980s, the housing turnover rate has risen steadily for the last 15 years until 2005. The result of a rising housing turnover rate indicates that home owners are capable of moving around more than in the past. In the housing market boom era, it also indicated the height of speculation. When the housing boom came to an end, the housing turnover rate started to decrease. The movement of housing turnover after 2005 shows exactly the same directionality. Since the housing turnover rate is used as the base prepayment speed and could generate a significant tail risk of principal loss given the same default probability, it is especially crucial for a high default and slow prepayment environment like the current one (Fig. 47.28).

AQ9

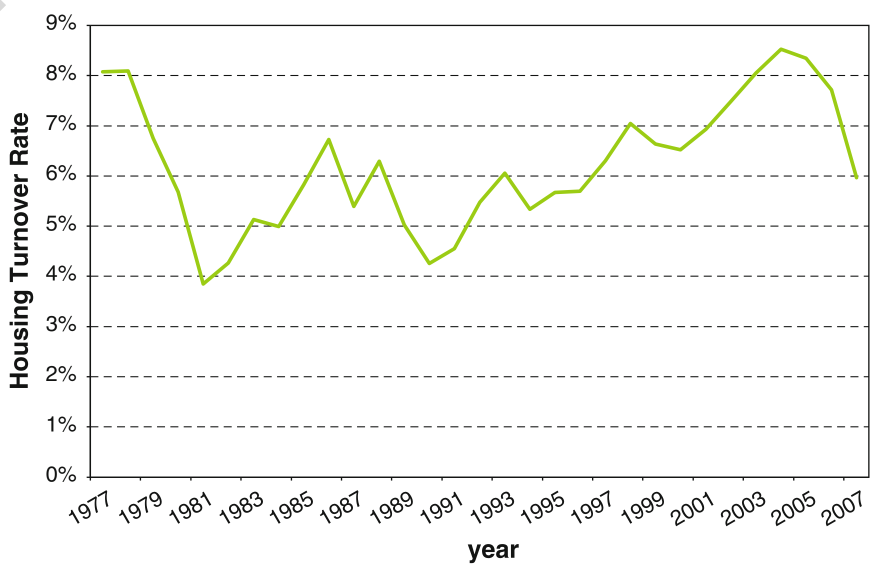
[13] We use a five-year moving average of "Total Occupied Housing Inventory" based on US Census Bureau times 0.67 to estimate the total single-family housing stock.

47.4.2 Seasoning

The initial origination fee and the loan closing expenses usually take a few years to be amortized, and this discourages the new mortgagors from prepaying their mortgages early in the mortgage term. This ramping-up effect is the seasoning factor. Figure 47.29 shows the age pattern observed for fixed rate loans. The ramping-up period initially lasts for the first few months and then it starts to level off or decrease due to other prepayment factors.

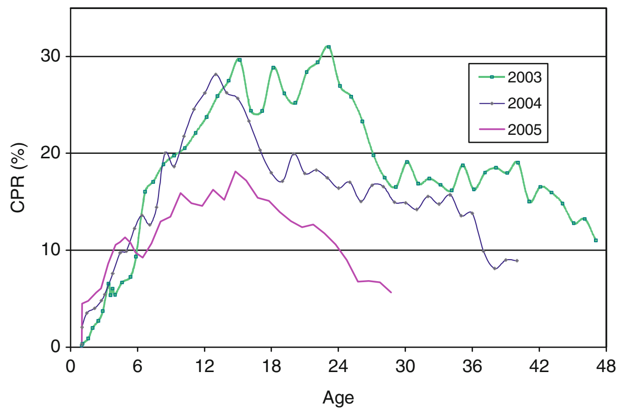
Hybrid ARMs exhibit similar patterns initially during the first 12 months. For hybrids like a ARM 2/28, the prepayment level climbs up from 0% to about 20–50% CPR within the first 12 months. After that, the acceleration of the prepayment levels starts to slow down until right before the teaser period ends. The difference in prepayment levels can be readily observed after the 12th month when shorter hybrids begin to show higher prepayment rates. The reason why ARM 2/28 borrowers show higher prepayment levels may be due to the faster housing turnover of the hybrid group. After the first 12 months, the prepayment generally stays around the same level with a wave-like trend peaking every 12 months. The seasoning pattern is illustrated in Fig. 47.30.[14]

[14] For the data pooling in terms of its vintage year, we usually use the loan distribution data for grouping information. It helps to maintain the relationship while examining the relationship with macroeconomic variable for time series analysis. It, however, distorts the age pattern since the loans within same vintage year could be underwritten in different months. The seasoning graph is specifically grouped by the loan's seasoning age to better understand the age pattern.



Sources: National Association of Realtors and Beyondbond

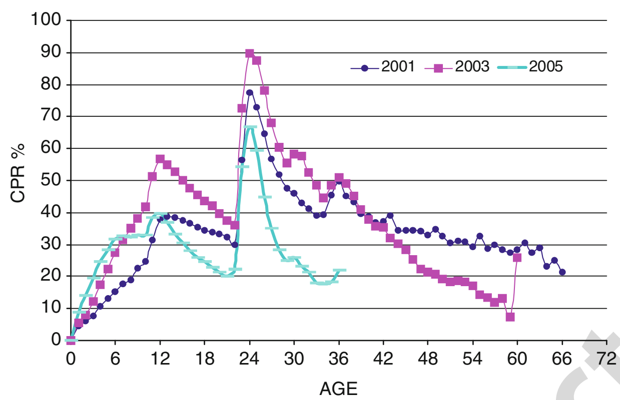
Fig. 47.28 US housing turnover 1977–2007 (Sources: National Association of Realtors and Beyondbond)



Sources: Beyondbond, Loan performance

Fig. 47.29 CPR over various vintages of discount fixed rate, coupon = 6% (Sources: Beyondbond, LoanPerformance)

B/W in Print



Sources: Beyondbond, Loan performance

Fig. 47.30 CPR over various vintages of ARM2/28 (Sources: Beyondbond, LoanPerformance)

B/W in Print

47.4.3 Teaser Effect

The teaser effect is the most distinctive feature of Hybrid ARM products. We define the term as the behavior that tends to persist right around the first reset where borrowers seek alternatives to refinance their mortgages or simply prepay them to avoid higher interest rates. In the following section we will describe the empirical statistics gathered to support the teaser effect.

Approximately 1 to 2 months before the end of the teaser period, a sharp rise in prepayments occurs. The effect is apparently larger for shorter hybrids like ARM 2/28 since shorter hybrids are exposed less to other prepayment factors such as refinancing and burnout before the teaser period. The peak level is reached just about 2 months after the teaser period ends. Teaser impact is usually observed as a sudden jump in prepayment levels. This spike happens whenever borrowers are able to refinance with a lower cost alternative.

47.4.4 Interest Only (IO) Effect

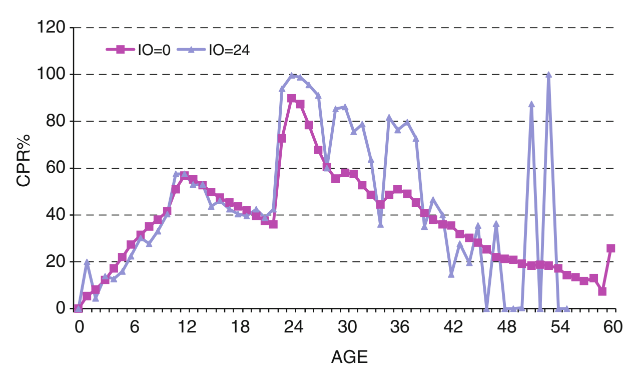
During the teaser period, the prepayments of ARM 2/28 with or without IO track each other fairly well. Before the end of the teaser period, loans with IO exhibit higher prepayment levels than the regular ones. IO borrowers are even more sensitive to the payment level since they are paying only the interest portion before the teaser. Once the teaser period ends, they will start to pay not only higher interest but also an additional amount of amortized principal. Their incentive to refinance is definitely higher than regular ARM 2/28 borrowers. Even worse, if they cannot find a refinancing alternative, they could face affordability issues and increased default risk (Fig. 47.31). We will address this further in the AQ10 interaction between prepayment and default section.

47.4.5 Refinance

The prepayment incentive is measured as the difference between the existing mortgage rate and the prevailing refinancing rate, which is commonly referred to as the refinance factor. As the refinancing factor increases, the financial incentive to refinance increases and thus changes prepayment behavior. When the loans are grouped by their coupon rates during the teaser period, the differences of prepayment levels are quite apparent. They behave in similar patterns, but loans with higher coupons tend to season faster due to the financial incentive to refinance, while loans with lower rates tend to be locked-in as the borrowers have secured the lower AQ11 rates (Fig. 47.32).s

47.4.6 Burnout Effect

The heterogeneity of the refinancing population causes mortgagors to respond differently to the same prepayment

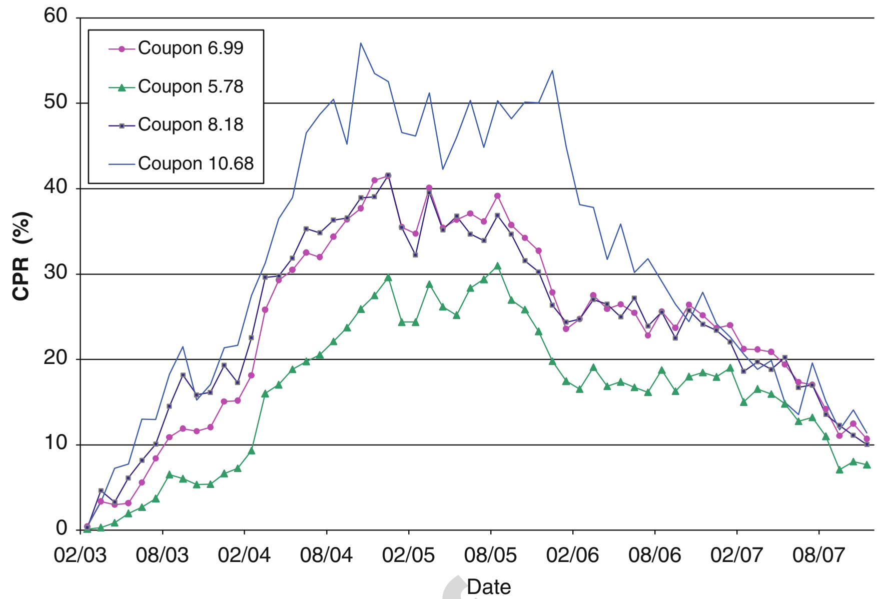


Sources: Beyondbond, Loan performance

Fig. 47.31 CPR over various vintages of ARM2/28 (Sources: Beyondbond, LoanPerformance)

B/W in Print

Fig. 47.32 Refinance: stratification by coupon, fixed rate, 2003 vintage (Sources: Beyondbond, LoanPerformance)



Sources: Beyondbond, Loan performance

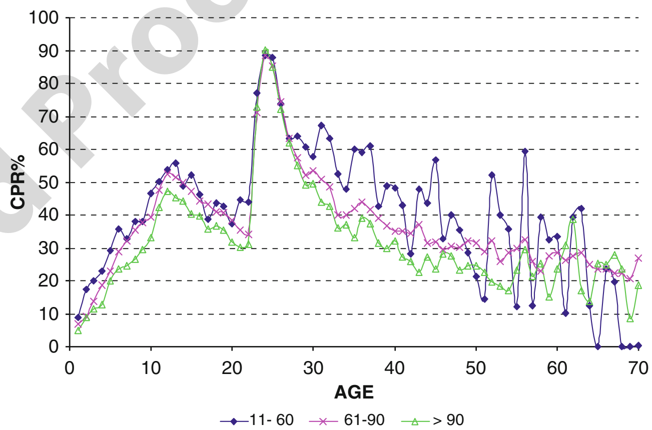
incentive and market refinancing rate. This phenomenon can be filtered out as burnout. The prepayment level usually goes up steadily with occasional exceptions across the high financial incentive region. The major reason for this is due to the burnout phenomenon in which borrowers that have already refinanced previously and have taken advantage of the lower rates are less likely to refinance again without additional financial incentives. To capture such a path-dependant attribute, our prepayment model utilizes the remaining principal factor to capture the burnout effect in order to reduce the chances of overestimating the overall prepayment levels.

47.4.7 CLTV Wealth Effect

As a property's price appreciates, the LTV of a loan gradually decreases. Borrowers with a low LTV may be able to refinance with a lower interest rate. Some borrowers may even find themselves in an in-the-money situation where they can sell their property to make an immediate profit. Historically, home prices continue to increase with age, and more and more loans will fall into this "low LTV" category, which has an increasing likelihood of prepayment. We use a combination of CLTV, HPA and age to model this effect (Figs. 47.33 and 47.34). AQ12

47.4.8 FICO Credit Effect

The subprime market consists of people with limited credit history and/or an impaired credit score. The high FICO score

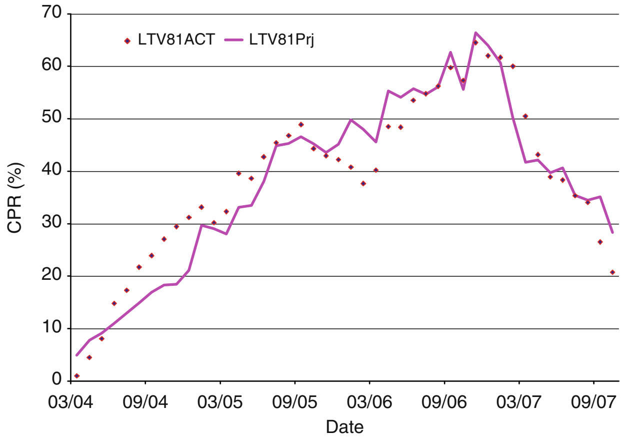


Sources: Beyondbond, Loan performance

Fig. 47.33 Wealth: CPR over various CLTV of ARM2/28 (Sources: Beyondbond, LoanPerformance)

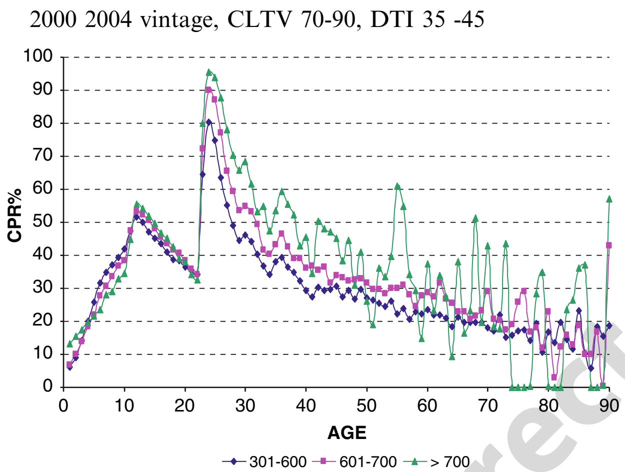
group is usually offered more alternatives to refinance and thus has the flexibility to choose between different products. For those people who are on the threshold of the subprime and prime market, they could be upgraded to participate in the prime market during the course of the loan life. Thus, the prepayment is an increasing monotonic function with respect to FICO (Fig. 47.35). AQ13

We can see a combined effect of FICO and CLTV on CPR. Those people who have a low CLTV and a high FICO score can easily refinance and will have the highest prepayment rate; while people who have high CLTV and low FICO score will be on the other side of the pendulum with the lowest prepayment rate. Figure 47.36 gives a sample CPR fitting result based on ARM 2/28, 2004 vintage pools.



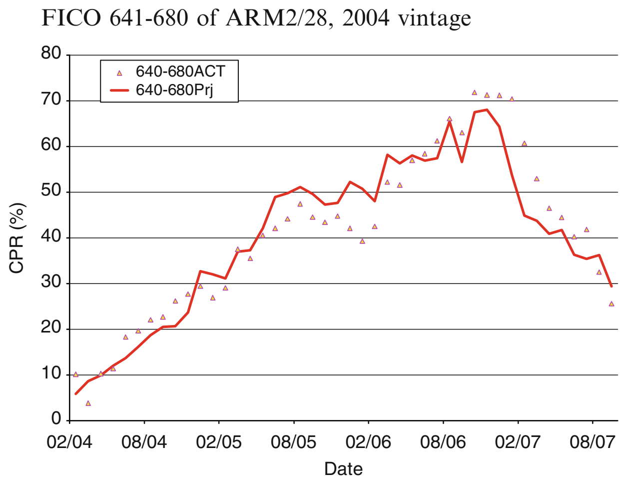
Sources: Beyondbond, Loan performance

Fig. 47.34 Fitted CPR over CLTV 81–90 of ARM2/28, 2004 vintage (Sources: Beyondbond, LoanPerformance)



Sources: Beyondbond, Loan performance

Fig. 47.35 Credit: CPR by FICO of ARM2/28c (Sources: Beyondbond, LoanPerformance)



Sources: Beyondbond, Loan performance

Fig. 47.36 Credit: fitted CPR (Sources: Beyondbond, LoanPerformance)

Sources: Beyondbond, Loan performance

Fig. 47.37 CPR over various vintages of ARM2/28 (Sources: Beyondbond, LoanPerformance)

47.4.9 Prepayment Penalty

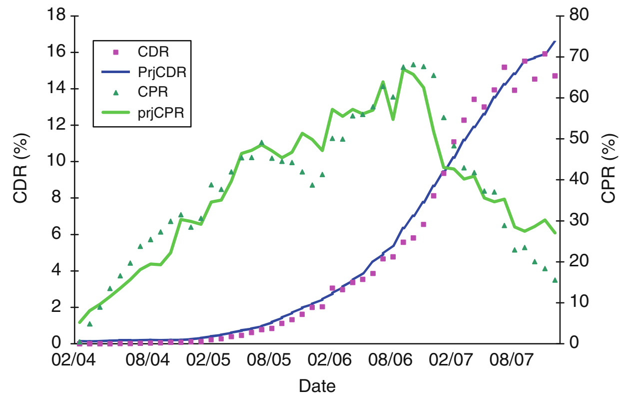
A prepayment penalty fee in the loan structure is a negative incentive and deters prepayment. Prepayment is essentially an embedded call option with the remaining balance as its strike. The penalty simply adds to that strike price as an additional cost when borrowers exercise the option. That additional cost will be reduced to zero when the penalty period ends. Figure 47.37 shows the prepayment difference when a penalty clause is in place. Before the 2-year penalty period ends, prepayment is consistently slower than no-penalty loans. As soon as the penalty period ends, prepayments surge dramatically and surpass the no-penalty loans within 3 months and consistently maintain a faster prepayment speed thereafter.

47.4.10 Interaction Between Prepayment and Default

As we stated in the beginning of the model framework, prepayment and default can be viewed as embedded call and put options, respectively, on the mortgage. A borrower will continuously find incentives to exercise it if the option is in-the-money (Fig. 47.38). AQ14

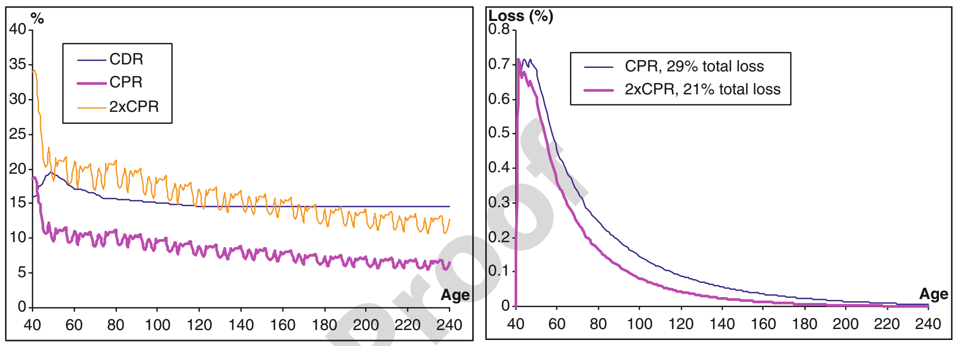
When we estimate prepayment and default for a pool of mortgages, the remaining principal factor encompasses the entire history of the pool's prepayment and default rates. Since estimating losses is the main focus for modeling default and prepayment, it is of particular importance in a slow prepayment environment. Given the same default probability, the tail risk to the loss curve will still increase

Fig. 47.38 CDR and CPR of ARM2/28, 2004 vintage (Sources: Beyondbond, LoanPerformance)



Sources: Beyondbond, Loan performance

Fig. 47.39 Loss projection of ARM2/28, 2004 vintage (Sources: Beyondbond, LoanPerformance)



Sources: Beyondbond, Loan performance

substantially. Figure 47.39 presents a tail risk example. When the prepayment speeds double, the total loss decreases to 21% from 29% given the same default speeds.

Because the history of prepayment and default rates can seriously affect the remaining principal factor for any given pool of loans, tracking and rolling the principal factor for a loan pool is one of the most important factors for the model projections and future forecasts. Prepayments are specified prior to defaults and are removed from the outstanding balance and, as a result, are not available to default in the future.

47.5 Delinquency Study

47.5.1 Delinquency, the Leading Indicator

Is delinquency a good leading indicator for default? When a borrower is late for a payment for more than 30 days, a 30-day delinquency is reported. If the payment is late for more than 2 months, a 60-day delinquency is reported. After a 90-day delinquency, the loan is considered to be in default, and the bank holding the mortgage will likely initiate its foreclosure process depending on the judicial status of each state. Since a default is a consequence of delinquency, the spectrum of delinquencies should be leading indicators of future defaults. We should be able to simply roll delinquency numbers month to month into actual defaults. The question is whether there is a constant relationship that can be parameterized or not. The time series plots of defaults and the spectrum of delinquencies for the 2003 vintage are shown in Fig. 47.40. The cross correlations indicate an approximately 6-month period for a 30-day delinquency to manifest into default as shown in Fig. 47.41.

47.5.2 Analysis Among Delinquency Spectrum

The results among delinquency spectrums show a very significant cross correlation between delinquency and its lagged earlier tenor (Fig. 47.42).

47.5.3 A Delinquency Error Correction Default Model

Based on the results shown previously, the spectrum of various delinquencies provides a good indication and can be parameterized for near-term projections. The benefit of

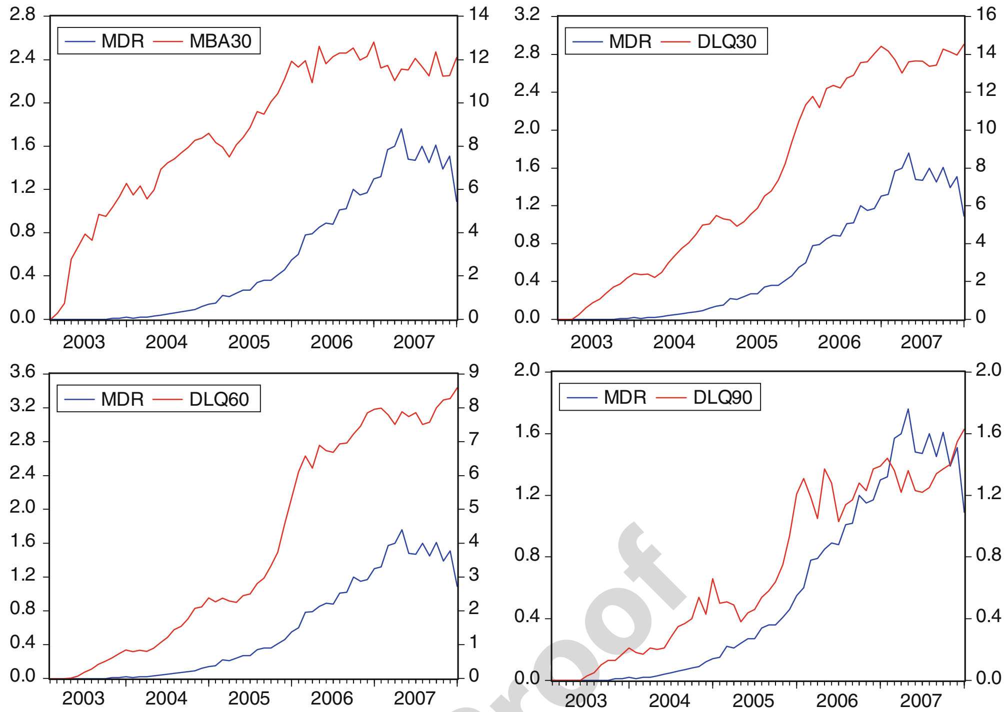


Fig. 47.40 Default and delinquency over time for 2003 vintage

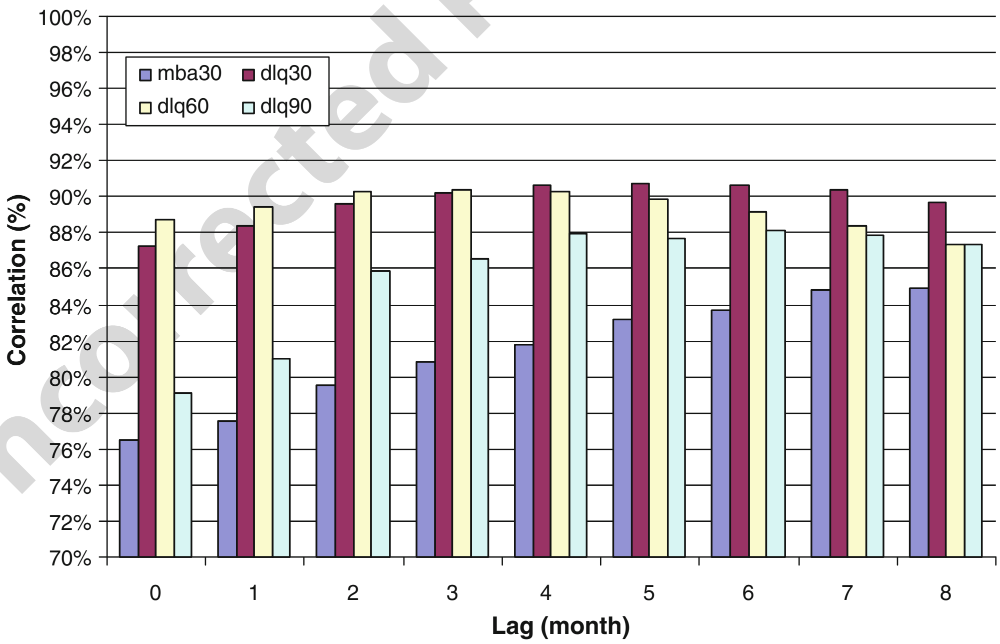


Fig. 47.41 Cross correlations of default and delinquency for 2000 vintages

including delinquency to project defaults is that it does not require specific consumer behavior theory to be applied. By simply looking at a delinquency report, we are able to project the likelihood of defaults. It, however, suffers from the long-term view that if a loan fundamentally carries lower credit-worthy characteristics such as a high CLTV, it has a greater propensity to default. Because we are impressed with their short-term forecast ability, and in order to fully utilize the information provided by delinquency and the econometric model based on consumer behavior theory, we have integrated both and created a delinquency error correction model.

	mba30	dlq30	dlq60	dlq90
mba30(-1)	0.974228	0.896283	0.849914	0.819303
dlq30(-1)	0.892006	0.99476	0.989324	0.931421
dlq60(-1)	0.842606	0.980814	0.993112	0.915923
dlq90(-1)	0.8199	0.937639	0.934675	0.898144

Fig. 47.42 Correlations between various delinquencies

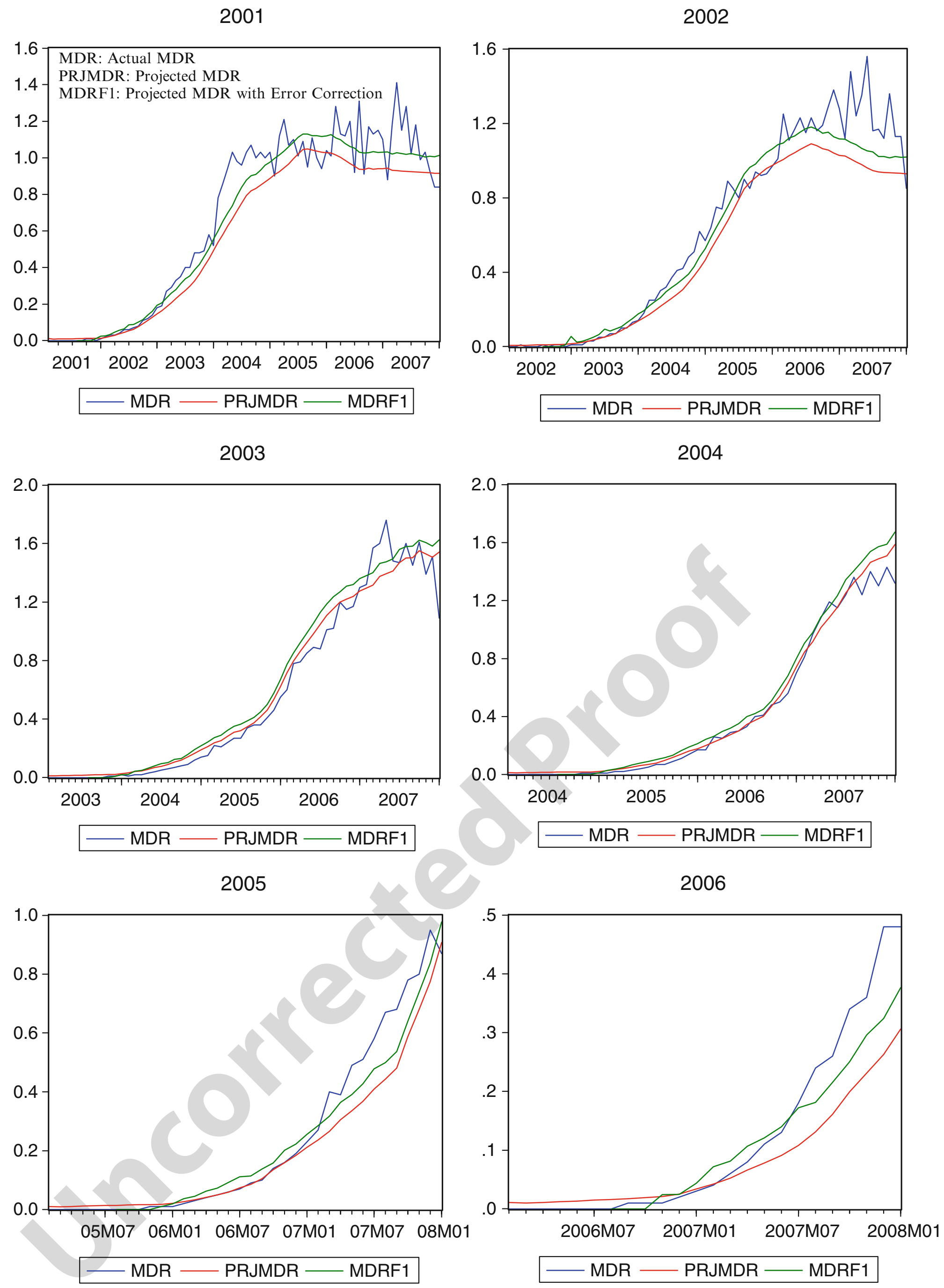
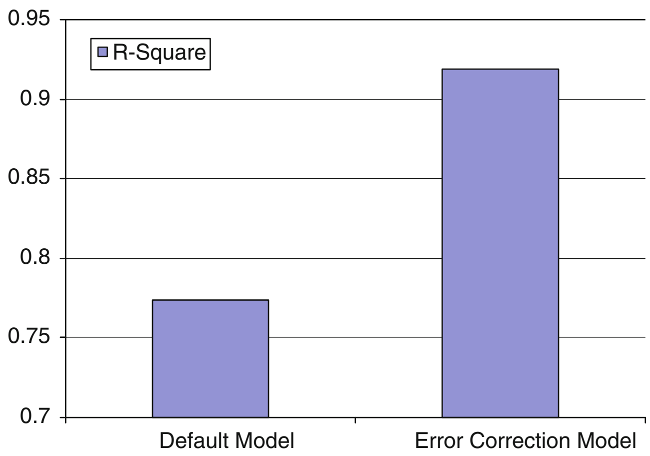


Fig. 47.43 Delinquency error model: actual vs. fitting

The fundamental idea is that not only can the long-term view and various scenarios based on changing views of macroeconomic variables be adopted, but also the immediate/early warning signs from delinquency can be observed and utilized (Fig. 47.43).

AQ17 In our error correction model, we start by projecting default rates using the default function with fitted parameters. We then layer on a 6-month lagged 30-day delinquency as an additional exogenous variable to regress the fitted errors. The process is then repeated sequentially by adding 5-month lagged 60-day and then 4-month 90-day delinquency rates as new regressors. The results are very encouraging when compared to the base model without error correction. The additional R^2 pickup is about 15% (Fig. 47.44). AQ18



Sources: Beyondbond, Loan performance

Fig. 47.44 Comparison of model explanation power for 2000–2007 vintages (Sources: Beyondbond, LoanPerformance)

47.6 Conclusion

47.6.1 Why Innovate?

As a result of the credit crisis, we now know we must have missed something in the traditional models. It requires us to take a hard look at the models and methodologies employed previously and see what is needed to provide a better interpretation of the current market data and conditions.

Traditionally, practitioners have observed consumer behavior through historical defaults and prepayments while building an econometric model with several quantifiable factors. These factors include seasoning patterns, underlying loan characteristics, such as mortgage coupon, FICO score, loan-to-value, and debt-to-income ratio, and macroeconomic variables, such as prevailing mortgage rate and housing price appreciation. In order to fit the historical data, non-linear functions are usually constructed with parameters around the factors to explain default and/or prepayment probabilities. During the process of historical sample fitting to the econometric model, the traditional modelers usually miss the following:

1. Traditional models focus on fitting in-sample data with a unique parameter set by vintage. Although the in-sample data fitting provides a much easier fit of the parameter set, it assumes that borrower's behavior varies given the same loan characteristics and loan age. It creates a disconnection among vintages and cannot be applied to new loans.
2. Borrower behaviors underlying LTV, FICO, and DTI were implicit but not fully quantified in a dynamic form by traditional models. Since the borrower and loan information such as LTV, FICO, and DTI levels are not periodically updated after the loan origination date, the accuracy of the projected performance of seasoned loans diminishes as time passes as the original data becomes aged and less relevant.
3. Out-of-sample projections may produce counterintuitive results. Macroeconomic variables, such as HPA, unemployment level, personal gross income, and so on can be very important factors for in-sample fitting. However, they do not provide insight for new scenarios. If a new scenario has not occurred historically, a stress test for the new scenario should be thoroughly pre-examined.
4. Traditional models focus on the national level rather than the local housing markets. Since house prices are highly dependent on location, a model with more detailed housing information can make a dramatic difference in the accuracy of its forecasts.
5. Traditional models treat prepayments and defaults independently and ignore the complexity and interaction between these embedded call and put options.
6. Traditional models do not dynamically quantify feedback from other leading indicators such as delinquency rates.

47.6.2 Innovation

Having addressed the pitfalls that traditional models fail to address, we have built a Dynamic Econometric Loss (DEL) model framework with the following innovations:

Consistent parameter sets for all vintages via the addition of consumer behavior factors.

1. Dynamic consumer behavior factors
   (a) CLTV ratio (via cumulative HPA since origination) that reflects housing market wealth effects during housing boom/bust eras.
   (b) DTI ratio (via unemployment rate forecasts) that addresses housing affordability.
2. Complete study of HPA index prior to model-fitting
   (a) HPCUM as the cumulative HPA since origination to capture wealth effect.
   (b) HPA to capture the pulse of the housing market.
   (c) HPA2D as the change of HPA to capture the trend of the housing market. HPA2D successfully captures the timing of defaults for 2005 to 2006 vintages.
   (d) In-sample and out-of-sample HPA fit testing to ensure the model's robustness.
3. A detailed CBSA-level HPA model allows us to understand local housing markets better and to generate more precise projections.
4. Recursive calculations along seasoning paths while estimating/projecting prepayments and defaults.

5. An error correction model that systematically builds the linkage between delinquency and default to enhance default forecast accuracy.

47.6.3 Advantages

The implementation based on our model framework will capture the default and loss patterns exhibited during the recent period and use the information contained in them to forecast future prepayments, defaults and losses based on various macroeconomic market scenarios. The implementation advantages are as follows:

1. Multiplicative and additive factors for each non-linear function (boot-strapping Maximum Likelihood Estimation)
2. Comprehensive consumer behavioral economic theory applied in practice
   (a) Develop a consumer behavior-based economic theory.
   (b) Estimate consumer behavior via an econometric model.
   (c) Apply the econometric model to prepayment and default.
3. Fully utilize HPA time-series information
   (a) A built-in time-series fitting model that dynamically estimates parameters and generates forecasts on the fly. For example,
      - HPCUM $\downarrow$(below 5%) $\Rightarrow$ CLTV$\uparrow\Rightarrow$ MDR$\uparrow$, SMM$\downarrow$
      - HPA $\downarrow$ (below 2%) $\Rightarrow$ MDR$\uparrow$, SMM$\downarrow$
      - HPA2D $\downarrow$(below -5%) $\Rightarrow$ MDR$\uparrow$, SMM $\downarrow$
4. Multiple built-in time-series fitting models at the national, state, and CBSA level that dynamically estimate parameters and generate forecasts on the fly.
5. Built-in recursive calculator along seasoning paths for projecting prepayments and defaults.
6. A set of error correction fitting models that estimate parameters within the spectrum of delinquencies and defaults.

47.6.4 Findings

In order to understand how a loan prepays or defaults, we have investigated consumer behavior via loan characteristics utilizing static factors and relevant macroeconomic variables as dynamic factors. For each factor, we have constructed a non-linear function with respect to the magnitude of the factor. We then built the default/prepayment function as a linear combination of these factors to justify the impact of each factor accordingly. Since a loan can either prepay or default over time, we then continue to ensure that the principal factors are rolled properly for prepayment and default forecasts.

While the level of HPA is considered the main blessing/curse for the rise and fall of the subprime market, we find that cumulative HPA and the change of HPA contribute additional dimensions to effect prepayment and defaults.

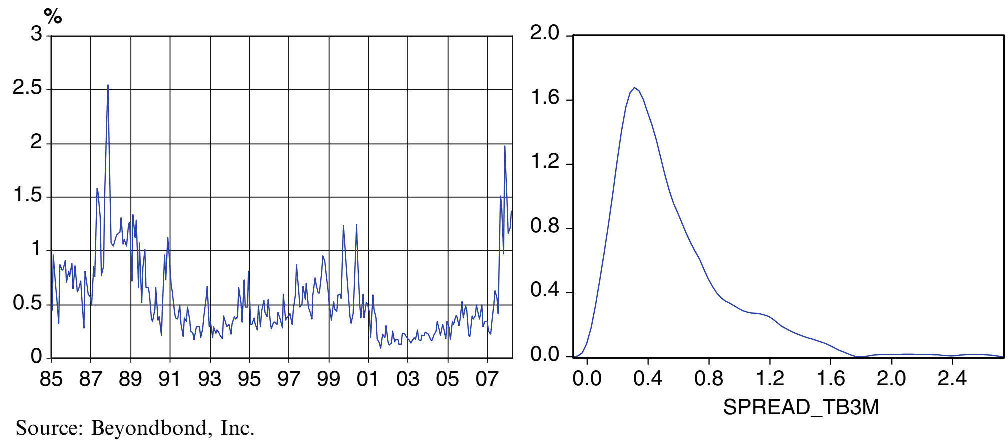
1. HPI is significantly correlated with DPI over a long-term period. Since DPI is a more stable time series, it suggests that HPI will eventually adjust to coincide with DPI growth rate.
2. Default is strongly correlated with the spectrum of delinquency rates. By applying the fitted parameters between default and delinquency rates to an error correction model, we are able to effectively improve default predictability.

47.6.5 Future Improvements

Modeling mortgage defaults and prepayments as embedded options is an ongoing learning process. While we are encouraged by our findings, there is a myriad of new questions for us to address with an aim to continuously improve and finetune the model. Some areas for further investigation are briefly described below.

47.6.5.1 Business Cycle – Low Frequency of Credit Spread

While studying the dynamic factors in the Default Modeling section, we focused mainly on the HPI impact on consumer behavior and introduced the DPI as another macroeconomic variable to determine the long-term growth of the economy. At the beginning of this paper, we wondered how a relatively small volume of loans could result in a subprime crisis that proved to be so detrimental to the entire US financial market and global financial system. We believe that the subprime crisis was merely the tipping point of unprecedented credit market easing that has existed since early this century. During this era of extremely easy credit, yield hungry investors sought to enhance their returns through investment in either highly leveraged securities or traditionally highly risky assets such as subprime loans. Through the rapid growth of the credit default swap in derivative markets and RMBS, ABS, and CDOs in the securitization markets, subprime mortgage origination volume reached record highs after 2003. The credit ease impacted not just the subprime market. All credit-based lending, from credit cards to auto loans and leveraged buy-out loans, were enjoying a borrower friendly environment as lenders went on a lending spree. While the credit default rates reached their historical low last decade and



Source: Beyondbond, Inc.

Fig. 47.45 Historical TED spread and histogram (Source: Beyondbond, Inc.)

resulted in extremely tight spreads among credit products, a longer view of the history of business cycles started to reveal warning signs of the potential downside risk.

For example, the TED Spread dramatically widened after August 2007, which was a re-occurrence of the late eighties market environment (Fig. 47.45). Over the past 20 years, traditional calibration models that only focused on shorter time frames missed the downside "fat tail." The improbable is indeed plausible. Is there a better method to mix the long-term low frequency data with the short-term high frequency data and provide a better valuation model?

AQ19

47.6.5.2 Dynamic Loss Severity

It is a usual practice, when using prepayment and default rates to forecast mortgage and mortgage-derived securities performance, to treat the lagged timing of loan loss/recovery and the loan loss/recovery level as given assumptions. The detailed HPA information provided at the CBSA-level and better detailed information from the loan servicers in recent years have allowed us to begin to model these variables to create dynamic loss severity percentages. Greater cooperation with the servicers will lead to more robust estimations.

References

AQ20

Bergantino, S. and G. Sinha. 2005. *Special report: subprime model update*, Bear Stearns & Co. Inc., May 26.

Dubitsky, R., J. Guo, L. Yang, R. Bhu, and S. Ivanov. 2006."Subprime prepayment, default and severity models." Credit Suisse, *Fixed Income Research*, May 17.

Flanagan, C. 2008. *Subprime mortgage prepayment and credit modeling*, J.P. Morgan Chase & Co., April 2.

Hayre, L. S. and M. Saraf. 2008. *A loss severity model for residential mortgages*, Citigroup Global Markets, Inc., U.S. Fixed Income Strategy & Analysis – Mortgages, January 22.

Hayre, L. S., M. Saraf, R. Young, and J. D. Chen. 2008. *Modeling of mortgage defaults*, Citigroup Global Markets, Inc., U.S. Fixed Income Strategy & Analysis – Mortgages, January 22.

Laderman, E. 2001. *Subprime mortgage lending and the capital markets*, Federal Reserve Bank of San Francisco, FRBSF Economic Letter (Number 2001–38), December 28.

Mago, A. 2007. *Subprime MBS: grappling with credit uncertainties*, Lehman Brothers Inc., March 22.

Mason, J. R. and J. Rosner. 2007. *Where did the risk go? How misapplied bond ratings cause mortgage backed securities and collateralized debt obligation market disruptions*, Mortgage-Backed Security Ratings, May 3.

Nera Economic Consulting. 2007. *At a glance: the chilling effects of the subprime meltdown*, Marsh & McLennan Companies, September.

Parulekar, R., U. Bishnoi, and T. Gang. 2008. *ABS & mortgage credit strategy cross sector relative value snapshot*, Citigroup Global Markets Inc., June 13.

Peterson, C. L. 2007. *Subprime mortgage market turmoil: examining the role of securitization – a hearing before the U.S. Senate Committee on Banking, Housing, and Urban Affairs Subcommittee on Securities, Insurance, and Investment*, University of Florida, April 17.

Ramsden, R., L. B. Appelbaum, R. Ramos, and L. Pitt. 2007. *The subprime issue: a global assessment of losses, contagion and strategic implications*, Goldman Sachs Group, Inc. – Global: Banks, November 20.

Risa, S. 2004. *The New Lehman HEL OAS model*, Lehman Brothers Inc., December 8.

Ted Hong, C. H. 2005. *Modeling fixed rate MBS prepayments*, Beyondbond Inc, October.

Ted Hong, C. H. and M. Chang. 2006. *Non-agency hybrid ARM prepayment model*, Beyondbond Inc, July, 6–17.

Wang, W. 2006. "Loss severity measurement and analysis." *The MarketPulse*, LoanPerformance, Issue 1, 2–19.

Appendix I Default and Prepayment Definition

We consider a loan to be in default if it meets both of the following criteria:

1. The loan is not able to generate any future investor cash flow

2. The loan has been in foreclosure, REO, or reporting loss in the prior reporting period

The Monthly Default Rate (MDR) is defined as the percentage of defaulted amount as a sum of all default loan balance compared with the aggregate loan balance of that period.

SMM (Single Month Mortality) is calculated by formula:

$$SMM = \frac{\text{Scheduled Balance} - \text{Current Balance}}{\text{Scheduled Balance}}$$

If we have MDR and SMM, then we can derive CDR and CPR from them by using the formula:

$$CDR = 1 - (1 - MDR)^{12}$$
$$CPR = 1 - (1 - SMM)^{12}$$

Appendix II General Model Framework

$$y_t^{(s)} = \sum_{k=0}^{K} \phi_k \left(X_t^{(k)} \left| \alpha_m^{(k)}, \beta_m^{(k)}; m \in [0, M^{(k)}] \right. \right) \cdot$$
$$\prod_{i=0}^{I} \lambda_i \left(X_t^{(i)} \left| \alpha_m^{(i)}, \beta_m^{(i)}; m \in [0, M^{(i)}] \right. \right) \cdot$$
$$\prod_{j=0}^{J} \eta_j \left(X_{t,m}^{(j)} \left| \alpha_m^{(j)}, \beta_m^{(j)}; m \in [0, M^{(j)}] \right. \right)$$
$$= \sum_{k=0}^{K} \phi_k \left(X_t^{(k)} \left| \alpha_m^{(k)}, \beta_m^{(k)}; m \in [0, M^{(k)}] \right. \right) \cdot$$
$$\prod_{i=0}^{I} \lambda_i \left(X_t^{(i)} \left| \alpha_m^{(i)}, \beta_m^{(i)}; m \in [0, M^{(i)}] \right. \right) \cdot$$
$$\prod_{j=0}^{J} \left(\left(1 + \sum_{m}^{M^{(j)}} X_{t,m}^{(j)} \beta_m^{(j)} \right) \left| \alpha_m^{(j)}, \beta_m^{(j)}; m \in [0, M^{(j)}] \right. \right)$$

where

- $y_t^{(s)}$ is an observable value at time t for dependent variable type s
- φ_k is a spline interpolation function with pair-wise $\left(\alpha_m^{(i)},\ \beta_m^{(i)}\right)$ knots
- $X_t^{(k)}$ is an observable value of factor k at time t
- K is the number of additive spline functions
- λ_i is a spline interpolation function with pair-wise $\left(\alpha_m^{(k)},\ \beta_m^{(k)}\right)$ knots
- $X_t^{(i)}$ is an observable value of factor i at time t
- I is the number of multiplicative spline functions
- η_j is equal to $\left(1 + \sum_m^{M^{(j)}} X_{t,m}^{(j)} \beta_m^{(j)}\right)$ and is a linear combination function with multiplier $\beta_m^{(j)}$ of $X_{t,m}^{(j)}$; where $X_{t,m}^{(j)}$ is an observable value of the type m factor at time t, while $\beta_m^{(j)}$ is the composition ratio of the distinct factor j of type m
- J is number of linear functions

Appendix III – Default Specification

A whole loan mortgage starts at t_0 and matures by t_n, its MDR by time t can be driven by two types of variables – static and dynamic.

Collateral characteristics such as mortgage rate, loan size, IO period, teaser period, loan structure, term to maturity, geographic location, FICO, and CLTV are static factors since their impact diminish over time while the loan is getting seasoned.

Macroeconomic variables over time such as Housing Price Index, mortgage interest rate, unemployment rates, Gross Disposable Income, and inflation rates are dynamic. They are publicly observable and will adjust the default rate forecasts based on the scenario assumption.

We formulate our default function MDR as follows:

$$D_t = \phi_{LTV}\left(v_t \left| LTV_j, h_t \right.\right) + \phi_{FICO}(c_j) \cdot$$
$$\lambda_{rate}(r_t|WAC_t) \ \cdot \lambda_{age}\left(a_i|a_0\right) \ \cdot \ \lambda DTI\left(\mathrm{d}_j|\mathrm{DTI}_j, \mathrm{DOC_j}\right) \ \cdot$$
$$\lambda_{IO}\left(g_t|IO_j, a_i\right) \ \cdot \lambda_{size}\left(s\right) \cdot \lambda_{HPA}\left(HPA\right) \cdot \lambda_{H2D}\left(H2D\right) \cdot$$
$$\eta_{DOC}\left(Doc_m\right) \cdot \eta_{LIEN}\left(LIEN_m\right) \cdot \eta_{PURPOSE}\left(PURPOSE_m\right) \cdot$$

where

- φ's are spline functions in MDR % and are additive to form a base value
- λ's are spline functions as multipliers for the MDR adjustments
- v_t: CLTV by time t where initial CLTV is assumed at time t_0
- r_t: Ratio spread of WAC_t over original WAC rate
- c_j: FICO score of loan j
- a_i: Age of loan j
- d_t: DTI
- g_i: Remaining IO period if IO exists and is positive
- l_j: Size of loan j
- φ_{LTV}: Original LTV level & HPA_t

$$v_t = v_t\left(v_0, h_t, z_j\right)$$

H_{t_i}: HPI at time t_i since origination date t_0

z_t: Geographic zip code j, e.g., $z_1 = z\,(CA) = 1.3$
$z_2 = z\,(OH) = 1.1$
$z_3 = z\,(MI) = 1.01$
$z_0 = z\,(Other) = 1$

the function form of v_t

$$v_t = \frac{v_0 \cdot H_{t_{(i-lag)}}}{H_{t_{(0-lag)}}}.z_j$$

h_t: the functional form of h_t as simple AR(2) model

$$h_t = \beta_0^h + \beta_1^h h_{t-1} + \beta_2^h h_{t-2} + \varepsilon_t$$

Where all the parameters can be independently regressed by h_t's time series data

z_j: the functional form of z_j is setup as a dummy variables

$z_j = \beta_j^z * z_{(j)}$ if j = "CA" and parameter β_j^z can be calibrated by default data by bootstrapping the value

f_t: is the actual principal factor and will be either observed for in-sample filtering or simulated for out-of-sample forecast

FICO: Checks if credit scores (original) are a good measure of default

c_j: the functional form of c_j will be a spline (natural, Linear, tension spline) function with

fixed FICO locators, j's (suggested only)
[250, 350, 450, 500, 525, 550, 550, 580,
600, 625, 650, 680, 700, 720, 750, 800, 820]

and parameters can be calibrated for default data base & fine-tuned

AGE: Default probability increases as loan get seasoned but eventually reach a plateau given other constants

a_t: we will sample linear spline function from 0 to 1 to apply age locators

[0, 1, 5, 10, 15, 20, 30, 45, 60, 120]

DTI Effect: Income level will affect default under assumption of DOC if it's fully available

$$u_t = u_0 \frac{GDP_t}{GDP_0} \cdot \left(\frac{UM_t}{UM_0}\right)^{\beta^{(UM)}}$$

the functional form

$\lambda_u\,(u_t)$ is a linear spline function of u_t
$\lambda_{DTI}\left(u_t, w_j\right) = (\lambda_u\,(u_t))^{\lambda_w(w_j)}$

where

$\lambda_w\,(w_0) = 1 \rightarrow \text{Full} = w_0$
$\lambda_w\,(w_1) = 0.1 \rightarrow \text{Low} = w_1$
$\lambda_w\,(w_2) = 0 \rightarrow \text{No} = w_2$

RATE Effect

$$r_t = (WAC_t - MTG_t)$$

$\varphi_{rate}\,(r_t)$ is a spline function of r_t

- WAC_t is gross coupon that is either observable or can be simulated from index rates and loan characteristics
- Index rate forecasting will be a spread

$$y_{t's} = \beta_0 + \beta_0 y_{t-1} + \beta_1 Swp2Y_t + \beta_2 Swp5Y_t$$
$$+\beta_3 Swp10Y_t + \beta_4 LIBOR1M_t + \varepsilon_t$$

for corresponding index rate LIBOR6M, 1Y-CMT, COFI, 5YY – CMT, . . .etc.

IO-Payment-Shock: Increased payments at the end of IO period will increase defaults.

$$g_t = IO_0 - a_t$$

$\lambda_{IO}(g_t)$ = is a linear spline function of locators [−30, −20, −10, −5, −2, 0, 2, 0, 2, 5, 10, 20]

Crowding Out: Measures if the underwriting standard is deteriorated

λ_{volume} is a spline function

vm_t is whole loan issue amount ratio (FICO $\leq$ 580, 580 $<$ FICO $\leq$ 700)

*Note: 30-day Delinquency rate for the (12-month) ratio if delinquency report is available

λ_{size} is a simple step-spline function to certain loan size after default with locators

[$\leq$ 50k, $\leq$ 100k, $\leq$ 150k, $\leq$ 250k, 500k, 800k, 1million]

Occupancy

λ_{ocp} has 3 kinds of occupancy (Owner, Second Home, Investor,)

Loan Purpose

λ_{prs} has 3 kinds of purpose (Purchase, Refi, Cash Out)

Lien

λ_{lien} has 2 lien positions (First lien, Second lien)

Loan Document

λ_{doc} has 3 kinds of documentation type (Full, Limit, and No Document)

Appendix IV – Prepayment Specification

Single Monthly Mortality (SMM) Rate Function

$S_t = \phi_{rate}\ (r_t)\cdot$
$\lambda_{turnoverrate}()\cdot$
$\lambda_{teaser}\ (ts_t)\cdot$
$\lambda_{seasonality}()\cdot$
$\lambda_{cash-out}()\cdot$
$\lambda_{age}\ (a_t)\cdot$
$\lambda_{burnout}(f_t)\cdot$
$\lambda_{yieldcurve}()\cdot$
$\lambda_{equity}()\cdot$
$\lambda_{credit}()\cdot$
$\lambda_{IO}\ (g_t)\cdot$
$\lambda_{credit}\ (V_t)\cdot$
$\lambda_{issuer}\ (IY_{j's})\cdot$
$\lambda_{size}\ (l_{j's})\cdot$
$\lambda_{penality}(N_{yes/no})$

Housing Turnover Rate

Prepayment based on long-term housing turnover rate and composed of existing sales over single-family owner housing stock.

Seasonality

Monthly seasonality is generally believed to affect prepayments. The belief stems from the mobility of mortgagors, time of housing construction, school year, and weather considerations. For a specific month of the year and ceteris paribus, prepayment rates are directly affected by the related month-of-year's coefficient. Usually, the seasonality pattern tends to be more active in the spring, rises to a peak in the summer, decreases through the fall, and slows down even more in the winter. The pattern may be different geographically and demographically.

Cash-Out

Prepayment is driven by general housing price appreciation.
Rate Factor $\varphi_{rate}(r_t)$ (to grab REFI-incentive)
φ_{rate}: a natural spline function
20 locators [−10, −5, −2, −1, 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7, 9, 10, 15, 20]

$$r_t = \begin{cases} WAC - m_t(\text{Fixed}) \\ WAC_D - m_t(\text{ARM/Hybrid}) \end{cases}$$

m_t: FH 30-yr/10 day commitment rate (FHR3010) as prevailing mortgage rate to measure SATO effect
*Age Factor: PPY has less incentive due to the consideration of initial financing sunk cost but the probability increases as 3-year costs average out over time.

Age

Mortgages generally display an age pattern.

Burnout Effect

Borrowers don't behave homogeneously when they encounter refinancing opportunities.

Some are more sensitive than others. If the borrowers are heterogeneous with respect to refinancing incentives, those who are more interest sensitive will refinance sooner. The remainder will be composed of less interest sensitive borrowers.

AUTHOR QUERIES

AQ1. Please provide Keywords.
AQ2. Citation for Figure 47.1 is inserted here. Please clarify if it is OK.
AQ3. Citation for Figure 47.2 is inserted here. Please clarify if it is OK.
AQ4. Citations for Figures 47.12 and 47.13 are inserted here. Please clarify if it is OK.
AQ5. Citation for Figure 47.20 is inserted here. Please clarify if it is OK.
AQ6. Citation for Figure 47.22 is inserted here. Please clarify if it is OK.
AQ7. Citation for Figure 47.23 is inserted here. Please clarify if it is OK.
AQ8. Citation for Figure 47.25 is inserted here. Please clarify if it is OK.
AQ9. Citation for Figure 47.28 is inserted here. Please clarify if it is OK.
AQ10. Citation for Figure 47.31 is inserted here. Please clarify if it is OK.
AQ11. Citation for Figure 47.32 is inserted here. Please clarify if it is OK.
AQ12. Citations for Figures 47.33 and 47.34 are inserted here. Please clarify if it is OK.
AQ13. Citation for Figure 47.35 is inserted here. Please clarify if it is OK.
AQ14. Citation for Figure 47.38 is inserted here. Please clarify if it is OK.
AQ15. Please check "...decreases to 21% from 29%..." and clarify if it is correct.
AQ16. Citation for Figure 47.42 is inserted here. Please clarify if it is OK.
AQ17. Citation for Figure 47.43 is inserted here. Please clarify if it is OK.
AQ18. Citation for Figure 47.44 is inserted here. Please clarify if it is OK.
AQ19. Citation for Figure 47.45 is inserted here. Please clarify if it is OK.
AQ20. All these references listed in the reference list are not cited in text. Please provide.