MODELING DEFAULTS IN BANKING & REAL ESTATE

by

Jiaqing Xu

B.S. in Applied Mathematics, Zhejiang University, 2007

M.A. in Applied Mathematics, University of Pittsburgh, 2009

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Jiaqing Xu

It was defended on

May 20, 2016

and approved by

Professor John Chadam, Department of Mathematics

Professor Brent Doiron, Department of Mathematics

Professor Song Yao, Department of Mathematics

Professor Satish Iyengar, Department of Statistics

Dissertation Director: Professor John Chadam, Department of Mathematics

MODELING DEFAULTS IN

BANKING & REAL ESTATE

Jiaqing Xu, PhD

University of Pittsburgh, 2016

In this thesis, we study two topics related to defaults. First, we provide a Probability of Default (PD) calculation method for privately-held U.S. regional banks, using free and transparent data from the Federal Deposit Insurance Corporation (FDIC). Our method is efficient and useful for both investors and regulators. We have improved Moody's proprietary RiskCalc PD model [17] by creating a new cautionary index, which is able to capture default behaviors very well and has a very high predictive power over both one-year and six-month time horizons as shown by our numerical results. We also find that this performance is robust over different historical periods. We describe the factors we chose, the modeling methodology, and the model's accuracy in detail.

Second, we propose two strategies to reduce the frequency of defaults in home mortgages (foreclosures). The first is a new mortgage insurance contract (American put option with the house as the underlying asset). Our analysis differs from that for the standard put option in equity markets in that our strike (the remaining value of the mortgage) is time dependent, and the drift and volatility in the Geometric Brownian Motion are time dependent (step functions) due to a regime switch from declining to increasing house prices. Both theoretical derivations and numerical results will be obtained. We will also analyze the Adjustable Balance Mortgage (ABM) in continuous time as a second alternative to avoiding foreclosures. Here the mortgage payments are reduced if the house price falls below the remaining value of the mortgage.

Keywords: Logistic Regression, Model Selection, ROC Curve, AUC, KS Statistic, American Put Option, Integral Equation, Early Exercise Boundary, Default Probability, Adjustable Balance Mortgage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First and foremost, I would like to express my special appreciation and thanks to my advisor Professor John Chadam - you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a researcher. Your advice on both research as well as on my career have been priceless. I would also like to thank my committee members, Professors Satish Iyengar, Brent Doiron, and Song Yao for serving on my dissertation committee. I also want to thank you for letting my defense be an enjoyable moment, and for your comments and suggestions - thanks to all of you.

A special thanks to my family. Words cannot express how grateful I am to my mother-in-law, father-in-law, my mother, father, and grandparents for all of the sacrifices that you've made on my behalf. Your prayers for me were what sustained me thus far. I would also like to thank all of my friends who supported me in writing, and encouraged me to strive towards my goal. And finally, I would like express appreciation to my beloved wife Zhuyun Xia who spent sleepless nights with and was always my support in the moments when there was no one to answer my queries.

1.0 INTRODUCTION

In the first part of the thesis, we provide a Probability of Default (PD) calculation method for privately-held U.S. regional banks, using free and transparent data from the FDIC. We cleaned the data, and applied a bidirectional stepwise logistic regression model to choose the best predictors out of eight predictors at each quarter t. Based on the corresponding estimated regression coefficients among all quarters, we choose the most important predictors. The final model for the PD estimation for all the banks requires only four predictors, and we do stepwise logistic regression on these four predictors. In order to check the predictive power, we apply the estimated coefficients from quarter t to quarter $t+1$. We then assess the goodness-of-fit of the model by three summary measures at each quarter: Area Under the ROC Curve, KS Statistic, and Bad Capture Rate in the Bottom Percentile. The results are shown to be very good for all quarters. We repeat this at each quarter and for different default horizons (1 year and 6 months). The ultimate cautionary index we create is a linear combination of the four predictors, where the weights are the actual coefficients of the corresponding significant predictors from the stepwise logistic regression deriving from the four significant predictors at each quarter. At time t, we set a subjective barrier (or cut-off level) based on the sorted cautionary index. We pick the barrier to be the index value such that the Bad Capture Rate Below The Barrier is at least 0.9 at time t (note that banks below the barrier, i.e., those with index value greater than the barrier, are those banks with high PD). We then carry over the coefficients as well as the barrier, and use them to predict 1 year in the future, out of sample, which defaults result; i.e. we apply the barrier from time t to time $t+1$, and check the Bad Capture Rate Below The Barrier at time $t+1$. The results prove to be very good. Most of the Bad Capture Rate Below The Barrier at time $t+1$ are greater than 0.9 for both 1 year and 6 months time horizons.

In summary, we have created a new cautionary index that improves on the proprietary Moody's model. It is capable of capturing the default behavior of risky banks with high predictive power. It should be quite useful to both the FDIC that regulates regional banks and for investors.

In the second part of the thesis we study two strategies to avoid defaults of home mortgages (foreclosures). The first is a new mortgage insurance contract, offering the mortgage holder an option to exchange the house for the remaining value of the mortgage, with no credit consequences. Essentially this new financial instrument is an American put option with the house as the underlying asset and the remaining value of the mortgage as the strike. The new technical features that arise are that the strike in the intrinsic value of the put option and the drift and volatility in the Geometric Brownian Motion are time dependent. The early exercise boundary, the value of the American put option, and the survival probabilities are calculated in both non-regime switching (falling house prices) and regime switching (falling followed by rising house prices) scenarios. We also study adjustable balance mortgages in continuous time. This second strategy allows the homeowner to pay a reduced mortgage payment if the house price falls below the remaining value of the mortgage.

In summary, these new financial instruments minimize the need to use the legal foreclosure system to deal with the economic risk of house price declines. By reducing the frequency of foreclosure, they benefit both the mortgage holders and the mortgage lenders. The mortgage holder does not suffer the adverse future credit impact from a foreclosure and the mortgage lender is not forced to hold a non-performing loan and to acquire a devalued property.

2.0 DEFAULTS OF PRIVATE REGIONAL BANKS

2.1 BACKGROUND

Regional banks play a central role in providing loans to businesses and farms, mainly for real estate. As a result, when there is a catastrophe in the real estate market, this affects the banks. It not only affects the investors of the banks, but also results in difficulties for the FDIC that regulates these banks. Specifically, the FDIC must become involved in the receivership process, pay for the insured deposits, and attempt to find buyers of the defaulted banks. There are more than 6,000 community banks in the U.S., representing 93% of all lenders in the country, and accounting for 45% of all small loans to businesses and farms. Many privately-held U.S. regional banks failed during the recent financial crisis, causing problems for investors and regulators.

Moody's RiskCalc PD model methodology document [17], states that their RiskCalc for U.S. Banks performs better at predicting bank failures than other publicly available models. However, the method is not transparent and it is almost impossible to replicate Moody's results. Additionally, Moody's data is proprietary and it is very expensive to subscribe to their service to get PD estimations.

The objective in this first part of the dissertation is to provide a PD calculation method for privately-held U.S. regional banks, using free and transparent data, which is efficient and accurate for both investors and regulators, and has a better predictive power than Moody's method. An immediate problem is that private banks provide very little information relative to publicly traded banks for which huge amounts of equity and bond information is available from the market. For private banks, the only information which is available is that found in their quarterly reports to the FDIC. This will be used to calculate the predictors, which, in

turn, will be used in the logistic regression model to develop our new cautionary index.
In summary, in our model only free and transparent data is used. In addition, we have improved Moody's model by creating a new cautionary index, capable of capturing the default behavior with high predictive power.

2.2 LITERATURE REVIEW

2.2.1 Previous Related Work On Bank Default Modeling

Bank failures have stronger adverse effects on economic activity than other business failures as pointed out by Gilbert, Meyer and Vaughan [11]. The spill over effect of bank panic or systemic risk has a multiplier effect on all banks and financial institutions leading to a greater effect of bank failure in the economy. As a result, banking institutions are typically subjected to rigorous regulation, and bank failures are of major public policy concern in countries across the world. During 1984-1993, there was a significant increase in the number of individual bank failures and acquisitions in the United States. Also, the 2008 financial crisis led to the failure of a large number of banks in the United States. The FDIC closed 465 failed banks from 2008 to 2012. Therefore, quantitative models now become more and more important since it not only helps the regulators to narrow the scope in the detection of financially weak institutions, but also it assists investors to identify bad investments.
The literature on forecasting bankruptcy and firm failure dates back to the 1960s. Altman [1] provides a summary through the early 1990s. The vast majority of these studies rely upon discriminant analysis, probit/logit models, and time-varying hazard analysis (or its variants). Bovenzi, Marino and McFadden [3] estimate PD for U.S. federally insured commercial banks by applying probit statistical analysis. Three models were developed based on different selected variables, where the variables account for the inherent risks in the commercial bank lines of business (i.e., credit risk, portfolio diversification, internal controls, operative inefficiencies, capitalization, and interest rates). Each model outputs classification accuracy at various threshold levels. Gilbert, Meyer and Vaughan [10] built a logit model to

estimate PD within the next two years for Fed-supervised banks, using a set of financial ratios focusing on capital adequacy, asset quality, management competence, earnings strength, liquidity risk and market risk. Eventually, fourteen financial ratios were defined as relevant in predicting PD. The ratios are: total equity to total assets, non-performing loans to total loans, consumer loans to total assets, other real estate owned to total loans, non interest expense to total revenue, insider loans to total assets, occupancy expense to average assets, return on assets, interest income accrued to total loans, liquid assets to total assets, large time deposits to total assets, core deposits to total assets, natural logarithm of total assets, and total assets to total assets in the parent-holding company.

2.2.2 Contribution

Our PD calculation method for privately-held U.S. regional banks differs from these previous studies in a number of directions. First, only free and transparent data from the FDIC are used. Second, we improved Moody's proprietary RiskCalc PD model by creating a new cautionary index, which is able to capture default behaviors very well and has a very high predictive power over both one-year and six-month time horizon. Third, our model performance is robust over different historical periods.

2.3 MODELING AND ANALYSIS

2.3.1 Input Data

Since we focus on privately-held U.S. regional banks, their quarterly financial statements are the only information available to us. We downloaded raw data files from the FDIC's database. There are 62 "*.csv*" files at each quarter, and there are 46 quarters, so in total we have 2,852 files. Figure 1 shows a list of the reports during the quarter ending on 9/30/2008.

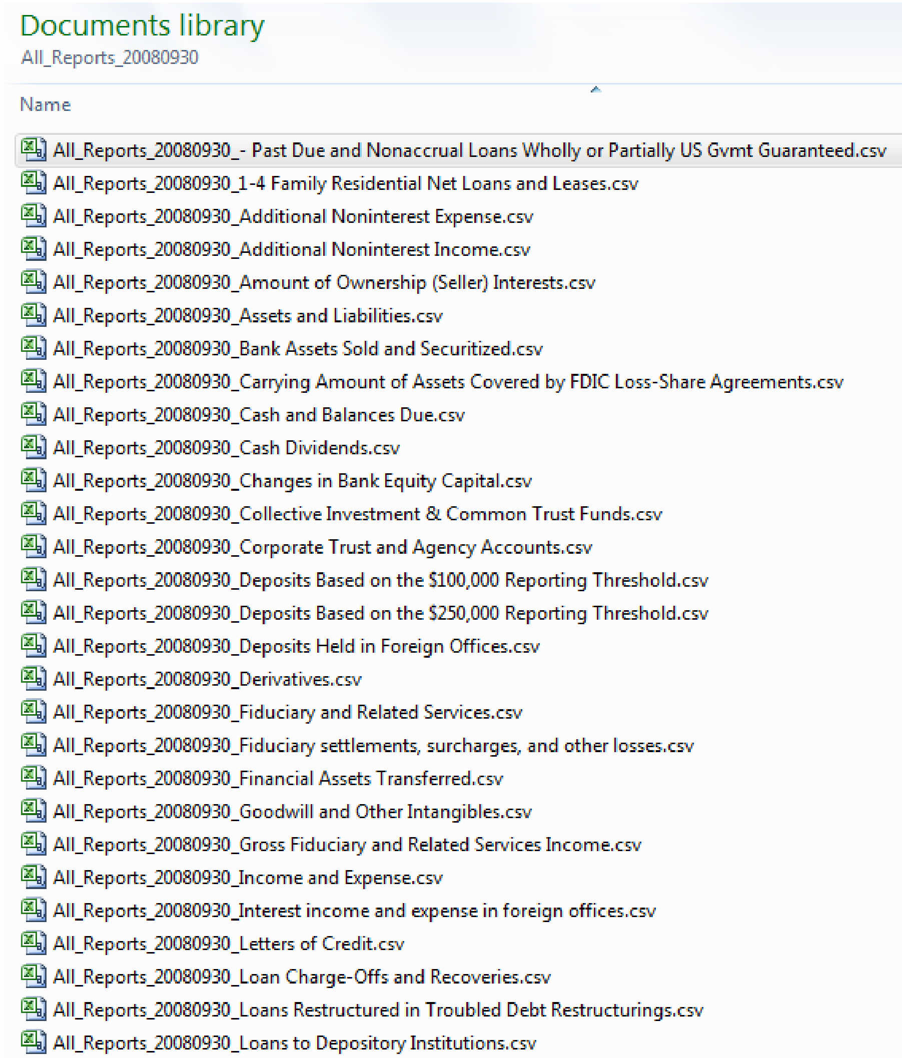

Figure 1: Sample reports on 9/30/2008

As an example, "`All_Reports_20080930_Assets and Liabilities.csv`" is an 8,393 by 89 matrix, whose snapshot is in Figure 2. It provides a great deal of information such as the bank ID, address, as well as the quantitative numbers of interest in this analysis.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	cert	docket	fed_rssd	rssdhcr	name	city	stalp	zip	repdte	rundate	bkclass	address	namehcr	offdom	offfor	stmult	specgrp	subchaps	county	sa_metro	tro_name
2	35260	0	2854113	0	09Republi	Bountiful	UT	84010	9/30/2008	8/20/2012	NM	1560 Renaissance To		1	0	0	8	1	Davis	36260	Ogden-Cl
3	57899	0	3317192	3370517	1st Advan	Saint Pete	MO	63376	9/30/2008	8/20/2012	NM	240 Salt Li	1ST ADVA	1	0	0	4	0	Saint Char	41180	St. Louis, I
4	15448	0	1006858	1495818	1st Ameri	Hancock	MN	56244	9/30/2008	8/20/2012	NM	Main Stre	WEST CEN	2	0	0	4	0	Stevens		
5	16522	0	888253	2003975	1ST BANK	Evanston	WY	82930	9/30/2008	8/20/2012	SM	1001 Main	GLACIER B	8	0	0	4	0	Uinta		
6	22039	0	564856	1128415	1st Bank	Sidney	MT	59270	9/30/2008	8/20/2012	SM	120 Secon	1ST UNITE	2	0	0	2	1	Richland		
7	16419	17301	419255	1055780	1st Bank 8	Broken Bc	OK	74728	9/30/2008	8/20/2012	SM	710 South	SOUTHEAS	5	0	0	8	1	Mccurtain		
8	30367	5236	148470	0	1st Bank o	Sea Isle Ci	NJ	8243	9/30/2008	8/20/2012	SA	4301 Landis Avenue		3	0	0	5	0	Cape May	36140	Ocean City
9	4788	0	488055	1055904	1st Bank o	Troy	KS	66087	9/30/2008	8/20/2012	NM	212 South	DON-CO. I	1	0	0	2	0	Doniphan	41140	St. Joseph
10	31826	7509	589671	1058165	1st Bank C	Claremore	OK	74017	9/30/2008	8/20/2012	SA	1698 S Lyn	MORRILL E	6	0	0	4	0	Rogers	46140	Tulsa, OK
11	57298	0	3048487	3559844	1st Bank Y	Yuma	AZ	85364	9/30/2008	8/20/2012	NM	2799 Soutl	WESTERN	3	0	0	4	0	Yuma	49740	Yuma, AZ
12	58485	0	3594797	0	1st Capita	Monterey	CA	93940	9/30/2008	8/20/2012	NM	5 Harris Court, Bldg N		3	0	0	4	0	Monterey	41500	Salinas, CA
13	33025	0	1463705	2847357	1st Center	Redlands	CA	92373	9/30/2008	8/20/2012	NM	218 East St	1ST CENTE	6	0	0	4	0	San Berna	40140	Riverside-
14	57657	0	3247598	3602236	1st Centur	Los Angel	CA	90067	9/30/2008	8/20/2012	N	1875 Cent	1ST CENTU	1	0	0	4	0	Los Angel	31100	Los Angel
15	35456	0	2920773	3118513	1st Coloni	Collingsw	NJ	8108	9/30/2008	8/20/2012	N	1040 Hadc	1ST COLOI	4	0	0	4	0	Camden	37980	Philadelpl
16	58358	0	3465383	1247334	1st Comm	North Las	NV	89031	9/30/2008	8/20/2012	NM	5135 Cami	CAPITOL E	1	0	0	4	0	Clark	29820	Las Vegas-
17	11706	12404	338945	1204083	1st Comm	Sherrard	IL	61281	9/30/2008	8/20/2012	NM	407 Third S	ILLINOIS H	2	0	0	2	1	Mercer	19340	Davenpor
18	27552	10064	1412619	2784920	1st Consti	Cranbury	NJ	8512	9/30/2008	8/20/2012	NM	2650 Rout	1ST CONS	11	0	0	4	0	Middlesex	35620	New York-
19	58321	0	3452123	0	1st Enterp	Los Angel	CA	90017	9/30/2008	8/20/2012	NM	818 W. 7th St., Suite		2	0	0	4	0	Los Angel	31100	Los Angel
20	34393	0	2531562	2531553	1st Equity	Skokie	IL	60076	9/30/2008	8/20/2012	NM	3956 West	FIRST EQU	1	0	0	4	1	Cook	16980	Chicago-Jo
21	57611	0	3203996	3203987	1st Equity	Buffalo Gr	IL	60089	9/30/2008	8/20/2012	NM	1330 Dunc	NORTHWE	1	0	0	4	1	Cook	16980	Chicago-Jo
22	1149	15588	50957	3105803	1st Financ	Overland	KS	66223	9/30/2008	8/20/2012	SM	11120 We:	FIRST CAP	4	0	1	4	0	Johnson	28140	Kansas Cit
23	1673	0	526452	0	1st Financ	Dakota Du	SD	57049	9/30/2008	8/20/2012	NM	331 N. Dakota Dunes		2	0	0	3	0	Union	43580	Sioux City
24	11338	14282	234645	1209109	1st Indepe	Louisville	KY	40220	9/30/2008	8/20/2012	NM	8620 Biggi	MAINSOU	8	0	1	4	0	Jefferson	31140	Louisville-
25	58548	0	3655081	0	1st Manat	Parrish	FL	34219	9/30/2008	8/20/2012	NM	12204 County Road 6		1	0	0	4	0	Manatee	35840	North Por
26	6646	8805	480723	0	1st Nation	Lebanon	OH	45036	9/30/2008	8/20/2012	N	730 East Main Street		12	0	0	4	1	Warren	17140	Cincinnati
27	3564	11706	737632	3437791	1st Nation	Homestea	FL	33030	9/30/2008	8/20/2012	N	1550 Nortl	HOMETOV	6	0	0	4	0	Miami-Da	33100	Miami-Foi
28	26977	10790	1000511	2421896	1st Nation	East Liver	OH	43920	9/30/2008	8/20/2012	N	16924 St. (	TRI-STATE	6	0	1	8	0	Columbiana		
29	35517	0	2912367	3487367	1st Pacific	San Diego	CA	92121	9/30/2008	8/20/2012	SM	9333 Gene	1ST PACIF	9	0	0	4	0	San Diego	41740	San Diego
30	57157	0	2998660	2998651	1st Regen	Andover	MN	55304	9/30/2008	8/20/2012	NM	1777 Bunk	ALLIANCE	2	0	0	4	1	Anoka	33460	Minneapo
31	57633	0	1018927	0	1st Securi	Mountlak	WA	98043	9/30/2008	8/20/2012	SB	6920 220th St Sw, Sui		13	0	0	4	0	Snohomis	42660	Seattle-Ta
32	9087	10250	991340	1199602	1st Source	South Ben	IN	46601	9/30/2008	8/20/2012	SM	100 North	1ST SOUR(	83	0	1	4	0	St Joseph	43780	South Ben

All_Reports_20080930_Assets and

Figure 2: Assets and liabilities .csv file on 9/30/2008 snapshot

Using Moody's research on bank PD modeling [17] as a guide, we initially constrained our approach to eight ratios specific to the banking industry, seven of which are directly from Moody's research: **Equity Capital/Assets**, **Commercial Real Estate Loans/Assets**, **Construction Loans/Assets**, **C&I Loans/Assets**, **Government Securities/Assets**, **Net Interest Margin**, **Net Income/Assets**. However, not all of them will turn out to be useful in our cautionary index. We will also require another called **Texas Ratio** (TR). In the beginning of this research, we thought the TR would be sufficient by itself to be the cautionary index. However, it turned out not to be as effective as we required. This accounted for the addition of some of the ratios from Moody's. "Equity Capital/Assets", "Construction Loans/Assets", "Net Income/Assets", along with "TR" ultimately developed from our analysis as the significant factors during the recent real estate crisis which caused the regional banks to default. The definition of these four factors are:

Equity Capital/Assets: $= (Total\ Assets - Liability)/Total\ Assets$. Assets are defined as anything that a business owns, has value and can be converted to cash. The equity capital is all of this money minus the liabilities, things which the bank must pay out.

Construction Loans/Assets: Construction Loans include loans for all property types under construction, as well as loans for land acquisition and development. Because this was a real estate crisis, the construction loans which are not performing can cause a great deal of harm to the bank. Our model identifies this ratio to be significant as well.

Net Income/Assets: A company's total profit as a percentage of total assets.

Texas Ratio: $= \frac{Non-performing\ Loans\ +\ Real\ Estate\ Owned}{Tangible\ Common\ Equity\ Capital\ +\ Loan\ Loss\ Reserves}$

The four components in TR are defined below:

Non-performing Loans: Loans that are delinquent.

Real Estate Owned: Property owned by a bank after an unsuccessful sale at a foreclosure auction.

Tangible Common Equity: a measure of a company's capital which can be used to evaluate its ability to deal with potential losses. This is an important factor in determining insolvency from any cause.

Loan Loss Reserves: money that the bank set aside in anticipation of some loans going bad. Thus if we are to have a cautionary index, with a high value indicating a high default probability, then TR should have a positive coefficient, Equity Capital/Assets should have a negative coefficient, Construction Loans/Assets should have a positive coefficient, and Net Income/Assets should have a negative coefficient. The 4 predictors will be selected from our statistical analysis as being significant, and therefore will be the ones used in our cautionary index.

From the FDIC's raw database, we have 9,873 U.S. regional banks' quarterly financial statements data from Q4 2002 to Q1 2014. However, not all of the data can be used in a consistent manner to do the statistical analysis. Some of them have missing data while others have data which have been corrupted, giving the wrong kinds of numbers. In the end, after data preprocessing and filtering steps, there are 6,081 healthy banks and 416 default banks, ranging from Q4 2002 to Q1 2014. We believe this is a good sample.

2.3.2 Multiple Logistic Regression

The previous section provides an idea of the data and the factors that will be used to develop our cautionary index. We will build the cautionary index using logistic regression.

2.3.2.1 Introduction

Let π denote the probability in favor of a binary event, and the probability against the event is therefore $1 - \pi$. The odds in favor of the event is defined as the ratio of the two, i.e. $odds = \frac{\pi}{1-\pi}$. The natural logarithm of the odds is called the logit function, i.e. $logit(\pi) = \ln(odds) = \ln(\frac{\pi}{1-\pi})$. Note that while the odds is strictly positive, the logit can take any real value. The logit function forms the basis for logistic regression. The model can be stated as follows:

Consider a collection of p predictors denoted by the vector $\mathbf{x}' = (x_1, x_2, ..., x_p)$. Let the conditional probability that the event will happen (a bank defaults) be denoted by $P(Y = 1|\mathbf{x}) = \pi(\mathbf{x})$. The logit of the multiple logistic regression model is given by the equation

$$logit(\pi(\mathbf{x})) = \ln\left(\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right) = g(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p \tag{2.3.1}$$

which implies that

$$E(Y|\mathbf{x}) = P(Y = 1|\mathbf{x}) = \pi(\mathbf{x}) = \frac{e^{g(\mathbf{x})}}{1 + e^{g(\mathbf{x})}} = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_p x_p}}{1 + e^{\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_p x_p}} \tag{2.3.2}$$

2.3.2.2 Fitting the Multiple Logistic Regression

Assume that we have a sample of n independent observations $(\mathbf{x_i}, y_i), i = 1, 2, ..., n$. Fitting the multiple logistic regression model requires that we obtain estimates of the vector $\boldsymbol{\beta}' = (\beta_0, \beta_1, ..., \beta_p)$. The estimation method we applied here is maximum likelihood. The likelihood function is

$$l(\boldsymbol{\beta}) = \prod_{n=1}^{n} \pi(\mathbf{x_i})^{y_i} \left[1 - \pi(\mathbf{x_i})\right]^{1-y_i} \tag{2.3.3}$$

Thus the log likelihood is

$$L(\boldsymbol{\beta}) = \ln\left[l(\boldsymbol{\beta})\right] = \sum_{i=1}^{n} \left\{y_i \ln\left[\pi(\mathbf{x_i})\right] + (1 - y_i)\ln\left[1 - \pi(\mathbf{x_i})\right]\right\} \tag{2.3.4}$$

The procedure here is to use the maximum likelihood method to find the coefficients of the predictors as well as the intercept. The $p + 1$ likelihood equations are obtained by differentiating the log likelihood function with respect to the $p + 1$ coefficients. There is no analytic solution for these equations, so solving them requires numerical methods and software packages. We use the statistical software R to get the maximum likelihood estimators $\hat{\boldsymbol{\beta}}' = \left(\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p\right)$.

2.3.3 Model Selection

We are motivated by wanting to balance goodness of fit and penalization for model complexity.

Akaike Information Criterion (AIC):

The smaller the value of AIC, the better the model.

$$\begin{aligned} AIC &= -2LogLikelihood + 2(p + 1) \\ &= -2\ln\left(L(\hat{\beta}_0, ...\hat{\beta}_p, \hat{\sigma}^2|Y)\right) + 2(p + 1) \end{aligned} \tag{2.3.5}$$

Bayes Information Criterion (BIC):

The smaller the value of BIC, the better the model.

$$\begin{aligned} BIC &= -2LogLikelihood + (p+1)\ln(n) \\ &= -2\ln\left(L(\hat{\beta}_0, ...\hat{\beta}_p, \hat{\sigma}^2|Y)\right) + (p+1)\ln(n) \end{aligned} \tag{2.3.6}$$

Since AIC performs well with large sample size, as is our case, we chose minimizing AIC as our criterion.

Moreover, there are many procedures for arriving at this optimal choice of the weighting factors.

Types of Selection Processes in Stepwise Regression:

Forward Selection

Start with no potential predictors; At each step, add the predictor for which the resulting model has lowest value of AIC; Stop when AIC begins to increase.

Backward Elimination

Start with all potential predictors; At each step, delete the predictor that results in the lowest value of AIC; Stop when AIC begins to increase.

Forward & Backward Selection (Bidirectional Elimination)

Do one step of forward selection; Do one step of backward elimination; Repeat until no predictors can be added or removed, and the resulting AIC is the smallest.

We choose the Forward & Backward Selection.

2.3.4 Evaluating Model Performance

After doing this, we will evaluate the model performance using some common metrics.

2.3.4.1 Area Under the ROC Curve (AUC)

The Receiver Operating Characteristic (ROC) curve is a plot which illustrates the performance of a binary classifier system as its discrimination threshold is varied. It is created by plotting the true positive rate (TPR) versus the false positive rate (FPR), at various threshold settings. TPR is also known as sensitivity, which defines how many correct positive results occur among all positive samples available during the test. FPR is one minus the specificity or true negative rate, which defines how many incorrect positive results occur among all negative samples available during the test.

$$sensitivity = TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \tag{2.3.7}$$

$$specificity = TNR = \frac{TN}{N} = \frac{TN}{FP + TN} = 1 - FPR \tag{2.3.8}$$

In our case,

TP: defaulted banks correctly predict as default

FP: healthy banks incorrectly predict as default

TN: healthy banks correctly predict as healthy

FN: defaulted banks incorrectly predict as healthy

To draw a ROC curve, only TPR and FPR are needed. The ROC curve plots parametrically TPR(T) versus FPR(T) with T as the varying threshold (decision) parameter (all possible cut-off points) and illustrates the performance of a binary classifier system as its discrimination threshold is varied. The ROC curve is also known as the trade-off curve because it shows the trade-off between goods and bads - the percentage of total bads that must be accepted in order to obtain a given percentage of total goods. The area under the ROC curve (AUC), which ranges from 0.5 to 1, provides a measure of the likelihood that an observation which has $Y = 1$ will have a higher $P(Y = 1)$ than an observation which has $Y = 0$. The idea is to have this curve as steep as possible and have this AUC being as close to 1 as possible, as a way of indicating our methods are performing well.

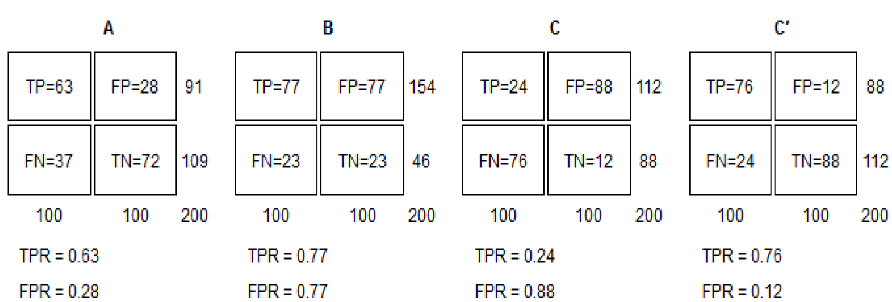

Figure 3: Four prediction examples.

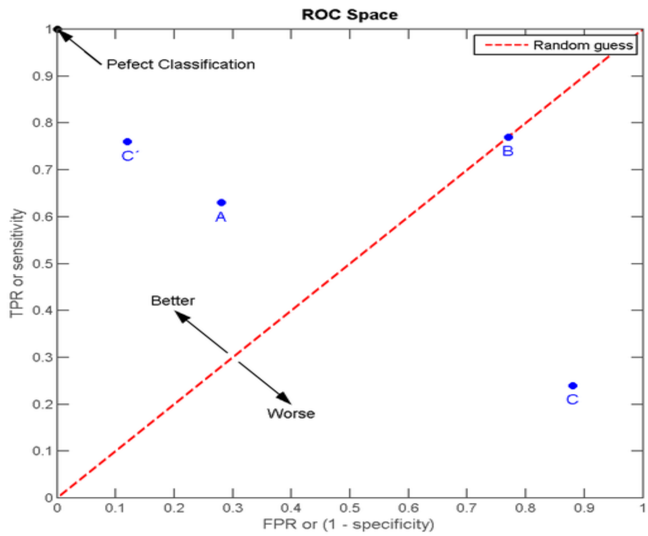

Figure 4: The ROC space and plots of the above four prediction examples.

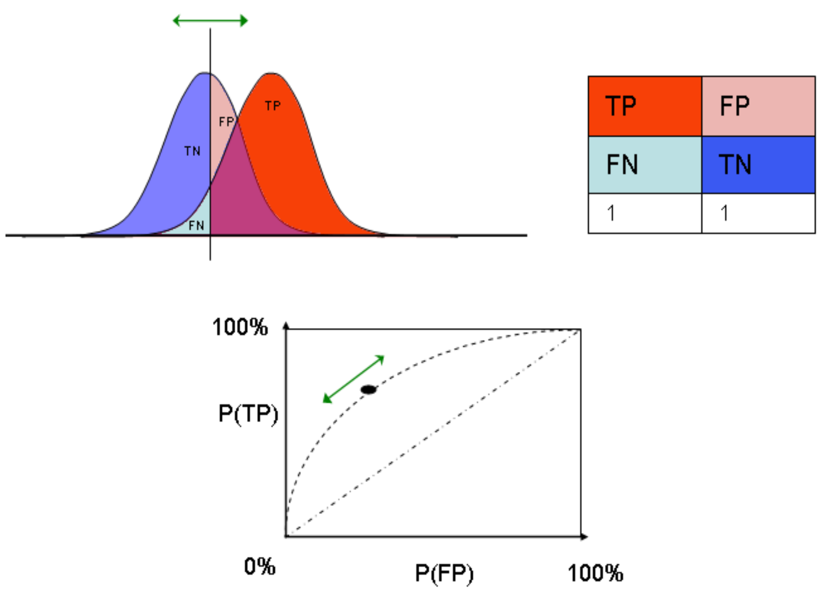

Figure 5: ROC curve intuition.

2.3.4.2 Kolmogorov-Smirnov(KS) Statistic (or KS measure)

The KS statistic is a widely accepted measure to evaluate the performance of a logistic regression. It gives the separation power a model exhibits between 0's and 1's. We want our model to assign higher PD to banks that eventually default than to banks that do not. The KS statistic is defined as the maximum difference between the cumulative percent good distribution and the cumulative percent bad distribution, i.e.

$F_G(PD)$ = CDF of the good observations

$F_B(PD)$ = CDF of the bad observations

$KS = \sup_{PD} |F_B(PD) - F_G(PD)|$

Theoretically, the KS statistic is a number ranging from 0 to 1. The higher the KS statistic, the better the separation power our model has.

A good graphical example of the K-S Statistic is in the credit lending business, which can be seen in Figure 6. The lending bank judges the credit quality of a person by the applicant's score which is derived from certain models. In this illustration, the greatest separability between the two distribution functions occurs at a score of approximately 0.7. Using this score, if all applicants who scored above 0.7 were accepted and all applicants scoring below 0.7 were rejected, then approximately 80% of all "good" applicants would be accepted, while only 35% of all "bad" applicants would be accepted. The measure of separability, or the K-S test result would be 45% (80%-35%).

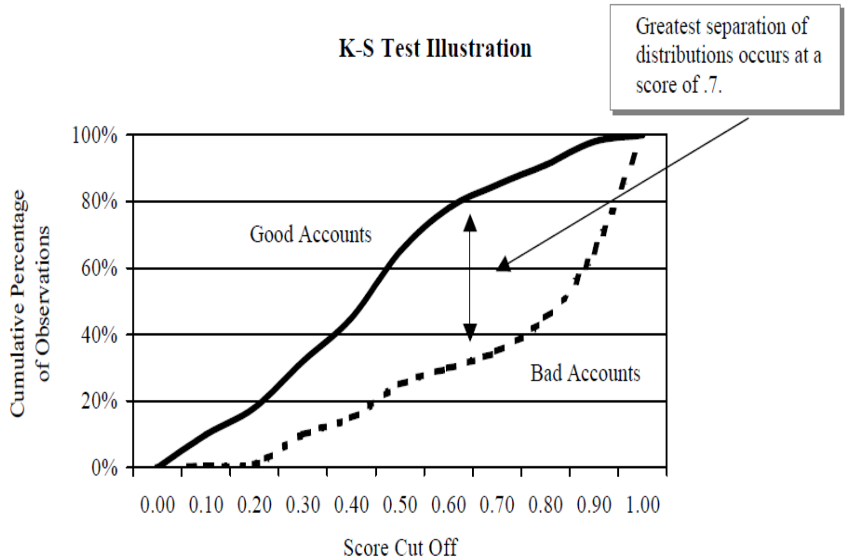

Figure 6: A graphical example of a KS Statistic.

2.3.4.3 Bad Capture Rate in the Bottom Percentile

The bad capture rate is a concept which is a more appropriate for the present analysis. What we are trying to do is to set some threshold above which we can capture the highest proportion of the banks which actually defaulted, without having too many false positives. This measure focuses on the bottom of the PD distribution (the highest PDs). The bad capture rate corresponding to the bottom percentage of all banks is defined as the number of default banks captured in the bottom percentile divided by the total number of defaulted banks. The higher the bad capture rate, the better the classification power of our model. We might choose the bottom 0.1%, 1%, 5%, 10%, 20%, 40%, or 50% as the cutoff. For example, if we have 6,000 banks in total at time t, then the bottom 1% will be the 60(=6,000*1%) banks whose estimated PDs from our model are the largest 1% among those banks. The bad capture rate corresponding to the bottom 1% is defined as the number of defaulted banks captured in the bottom 60 banks divided by the total number of defaulted banks.

2.3.4.4 How To Choose An Appropriate Bottom Percentile

As we increase the bottom percentile, the bad capture rate will increase and finally reach

1 (this happens after we capture all the default banks but we also have captured many successful banks). As mentioned previously, we want a higher bad capture rate, which is an increasing function of the bottom percentile. At the same time, as we increase the bottom percentile, the ratio between the number of healthy banks within the bottom and the number of defaulted banks within the bottom increases. This ratio is a measure of how many good banks must be observed in order to capture the bad banks. Of course, we want this ratio to be small. Therefore, there is a trade-off between the good over bad ratio and bad capture rate. So we must determine a rule to choose an appropriate bottom percentile.

2.4 RESULTS

In summary, we used cleaned FDIC data, performed bidirectional stepwise logistic regression using AIC as the criterion, and evaluated the model performance by the above three metrics. We apply the stepwise logistic regression model to choose the best predictors out of the eight predictors at each quarter t. Based on the corresponding estimated regression coefficients among all quarters, we choose the most important four predictors.

Our final model for the PD estimation for all the banks uses only these four predictors. We then do stepwise logistic regression on these four predictors.

Table 4 and Table 5 in the Appendix summarize the stepwise logistic regression coefficients in each quarter for 1 year default time horizon. Table 6 and Table 7 in the Appendix provide the same summary for 6 months default horizon.

Note that in the tables, the header β_0 corresponds to the intercept in the regression, and $\beta_1, ..., \beta_8$ correspond to the coefficients of Texas Ratio, Equity Capital/Assets, Commercial Real Estate Loans/Assets, Construction Loans/Assets, C&I Loans/Assets, Government Securities/Assets, Net Interest Margin, and Net Income/Assets.

In each quarter, regardless of the p-values, the intercept, β_0, is retained. For each factor, if the p-value is less than 0.05, it is retained as significant. A blank indicates that the p-value for this factor was larger than 0.05 and hence was insignificant in this quarter. One notices that TR, EC/A, CL/A, and NI/A are significant with higher frequency than the other factors.

Based on these results, we maintain that the four most important predictors for both 1 year and 6 months default time horizons are Texas Ratio, Equity Capital/Assets, Construction Loans/Assets, and Net Income/Assets.

Using only these factors, we recalculate the coefficients using stepwise logistic regression at each time t, using the information about the banks which have defaulted in a 1 year (or 6 months) time horizon. We then use the calculated coefficients 1 year (or 6 months) later to check their effectiveness, i.e. we apply the coefficients estimated from time t to predict at time $t + 1$. We then assess the goodness-of-fit of the model by the summary measures at time $t+1$. We find this to be very good for all quarters. We repeat this at each quarter and for different default horizons (1 year and 6 months).

Table 1: Logistic regression predictive power summary by the selected four predictors: apply coefficients estimated from 20080930 to 20090930, 1 year default horizon

Start Quarter t	20080930						
Default Horizon	1 year						
Apply to t+1	20090930						
Number of Defaults at t	90						
Number of Defaults at t+1	147						
Stepwise Logistic Regression AIC	471.86						
ROC Curve AUC At Time t+1	0.977						
KS Statistic At Time t+1	0.917						
Bottom Percentage	0.1%	1.0%	5.0%	10.0%	20.0%	40.0%	50.0%
Bad Capture Rate t	6.7%	44.4%	85.6%	90.0%	94.4%	96.7%	97.8%
Bad Capture Rate t+1	4.1%	38.1%	89.8%	97.3%	97.3%	98.0%	98.6%

Table 1 summarizes the detailed stepwise logistic regression predictive power, where we apply coefficients estimated from 9/30/2008 data to the data on 9/30/2009 using a 1 year default horizon. We found that the AUC is good and the KS statistic is very close to 1 as well. The most significant result is the bad capture rate. This is the bad capture rate in sample, at 5% cut off. When we use the coefficients from time t to $t+1$ (out of sample), we get essentially the same efficiency.

It is clear that to obtain a 98.6% bad capture rate at time $t+1$ requires observing 3,233 banks which is too expensive. On the other hand, by setting the bottom percentage at 5% (observing 323 banks) or 10% (observing 647 banks) we capture (identify) between 89.8% and 97.3% of the 147 banks that will fail one year later. The results for other quarters during the housing crisis are similar.

The new cautionary index we create is a linear combination of the 4 predictors, where the weights are the actual coefficients of the corresponding significant predictors from the stepwise logistic regression deriving from the four significant predictors at each quarter. At time t, we set a subjective barrier (or cut-off level) based on the sorted cautionary index. We pick the barrier to be the index value such that the Bad Capture Rate Below The Barrier is at least 0.9 at time t (note that banks below the barrier are those with index value greater than the barrier, which are those banks with high PD). Then we carry over the coefficients as well as the barrier, and use them to predict 1 year hence out of sample what the results are going to be, i.e. we apply the barrier from time t to time $t+1$, and check the Bad Capture Rate Below The Barrier at time $t+1$. Table 2 and Table 3 show the detailed results which prove to be very good: most of the Bad Capture Rate Below The Barrier at time $t+1$ are greater than 0.9 for both 1 year and 6 months time horizons. Table 2 provides a summary of the analysis. What is important here is that starting about Q3 2007, there are significant numbers of defaults. They continue for a considerable time, until sometime around 2012. One sees that during late 2007 and early 2008 the number of healthy banks divided by the number of default banks below the barrier is much too high. There are too many banks to monitor in order to find the banks which actually default. However as one continues to Q3 2009, one sees there are only 285 banks to be monitored to capture 75 banks that defaulted. So in the period when there are significant defaults due to the real estate collapse, the number of healthy banks divided by the number of defaults on the bottom percentile (which gives the indication of how good the method is) starts to decrease to around 2, as opposed to 32.8 in year 2007.

The same observations apply to the 6 months default horizon (Table 3). Again the ratio of decreases from 152.8 to approximately 2 during the housing crisis. One notices, in addition, that the ratio jumps back up to 13 in the third quarter of 2013, indicating that perhaps the housing crisis is over.

Table 2: 1 year default horizon summary by using barrier on the cautionary index

t	Apply to t+1	Barrier (Weighted Index) at t	# of Banks On The Bottom Percentile at t+1	# of Defaults On The Bottom Percentile at t+1	# of Healthy On The Bottom Percentile at t+1	# of Healthy/# of Defaults On The Bottom Percentile at t+1	# of Default at t+1	Bad Capture Rate On The Bottom Percentile at t+1
2007.9.30	2008.9.30	0.7424893	2911	86	2825	32.8488372	90	0.95555556
2007.12.31	2008.12.31	0.9900067	1403	116	1287	11.0948276	118	0.98305085
2008.3.31	2009.3.30	-2.657913	1072	131	941	7.18320611	135	0.97037037
2008.6.30	2009.6.30	-3.769981	1624	148	1476	9.97297297	154	0.96103896
2008.9.30	2009.9.30	-1.474294	773	143	630	4.40559441	147	0.97278912
2008.12.31	2009.12.31	-3.05916	427	131	296	2.25954198	134	0.97761194
2009.3.31	2010.3.31	-5.355394	439	111	328	2.95495495	117	0.94871795
2009.6.30	2010.6.30	-2.854472	340	90	250	2.77777778	94	0.95744681
2009.9.30	2010.9.30	-4.121945	285	75	210	2.8	78	0.96153846
2009.12.31	2010.12.31	-2.300902	171	62	109	1.75806452	66	0.93939394
2010.3.31	2011.3.31	-4.338152	196	59	137	2.3220339	61	0.96721311
2010.6.30	2011.6.30	-4.126456	200	57	143	2.50877193	59	0.96610169
2010.9.30	2011.9.30	-4.090674	155	49	106	2.16326531	51	0.96078431
2010.12.31	2011.12.31	-3.47596	131	44	87	1.97727273	45	0.97777778
2011.3.31	2012.3.31	-3.90407	97	33	64	1.93939394	35	0.94285714
2011.6.30	2012.6.30	-4.604535	102	27	75	2.77777778	28	0.96428571
2011.9.30	2012.9.30	-3.479197	57	21	36	1.71428571	23	0.91304348
2011.12.31	2012.12.31	-2.485187	47	16	31	1.9375	19	0.84210526
2012.3.31	2013.3.31	-4.590071	54	15	39	2.6	19	0.78947368

Table 3: 6 months default horizon summary by using barrier on the cautionary index

t	Apply to t+1	Barrier (Weighted Index) at t	# of Banks On The Bottom Percentile at t+1	# of Defaults On The Bottom Percentile at t+1	# of Healthy On The Bottom Percentile at t+1	# of Healthy/# of Defaults On The Bottom Percentile at t+1	# of Default at t+1	Bad Capture Rate On The Bottom Percentile at t+1
2007.9.30	2008.3.31	0.5291853	1384	9	1375	152.777778	9	1
2007.12.31	2008.6.30	0.5787903	688	16	672	42	16	1
2008.3.31	2008.9.30	-7.557068	1550	26	1524	58.6153846	28	0.92857143
2008.6.30	2008.12.31	-2.144762	585	38	547	14.3947368	39	0.97435897
2008.9.30	2009.3.31	-4.485205	1299	60	1239	20.65	62	0.96774194
2008.12.31	2009.6.30	-2.705297	259	70	189	2.7	79	0.88607595
2009.3.31	2009.9.30	-3.043753	230	66	164	2.48484848	73	0.90410959
2009.6.30	2009.12.31	-3.351872	472	73	399	5.46575342	75	0.97333333
2009.9.30	2010.3.31	-5.585763	273	72	201	2.79166667	74	0.97297297
2009.12.31	2010.6.30	-1.90663	105	57	48	0.84210526	59	0.96610169
2010.3.31	2010.9.30	-3.245775	109	39	70	1.79487179	43	0.90697674
2010.6.30	2010.12.31	-5.350716	98	31	67	2.16129032	35	0.88571429
2010.9.30	2011.3.31	-2.355462	67	30	37	1.23333333	35	0.85714286
2010.12.31	2011.6.30	-4.127371	117	30	87	2.9	31	0.96774194
2011.3.31	2011.9.30	-6.396689	84	24	60	2.5	26	0.92307692
2011.6.30	2011.12.31	-3.252554	53	24	29	1.20833333	28	0.85714286
2011.9.30	2012.3.31	-5.419467	74	23	51	2.2173913	25	0.92
2011.12.31	2012.6.30	-2.360364	18	13	5	0.38461538	17	0.76470588
2012.3.31	2012.9.30	-5.042031	32	10	22	2.2	10	1
2012.6.30	2012.12.31	-3.628586	49	11	38	3.45454545	11	1
2012.9.30	2013.3.31	-3.341269	9	7	2	0.28571429	13	0.53846154
2012.12.31	2013.6.30	-4.181472	9	4	5	1.25	8	0.5
2013.3.31	2013.9.30	-6.568359	84	6	78	13	6	1

3.0 DEFAULTS ON HOME MORTGAGES

3.1 BACKGROUND

The second problem of this dissertation concerns defaults in home mortgages, which are related to the collapse of the real estate market. The homeowner who defaults on his home mortgage later suffers credit consequences from the default.

In this work we explore two different ways to avoid mortgage defaults when housing prices decline. One way is to create a mortgage insurance that allows the buyer to stop paying the mortgages and walk away without any credit consequences. Another is to consider a different type of mortgage that allows the mortgage holders to pay less when the house prices decline. Firstly, we propose a new mortgage insurance contract (American put option with the house as the underlying asset). By reducing the role of the legal system in mitigating house price risk, this new financial instrument minimizes the need to use the legal foreclosure system to deal with the economic risk of house price declines. Secondly, we will price the adjustable balance mortgage in continuous time.

Historically, the housing market was on its way up from 2003 to 2006, and then there was a significant drop during the period 2006 and 2012. Data during this period was provided by Zillow, the real estate database company. In order to preserve the anonymity of the individual homeowners, they provided only the average house prices. Table 8 in the Appendix is Zillow's data from Paradise, Las Vegas. We will use the data to model the evolution of house prices during 1997 to 2013.

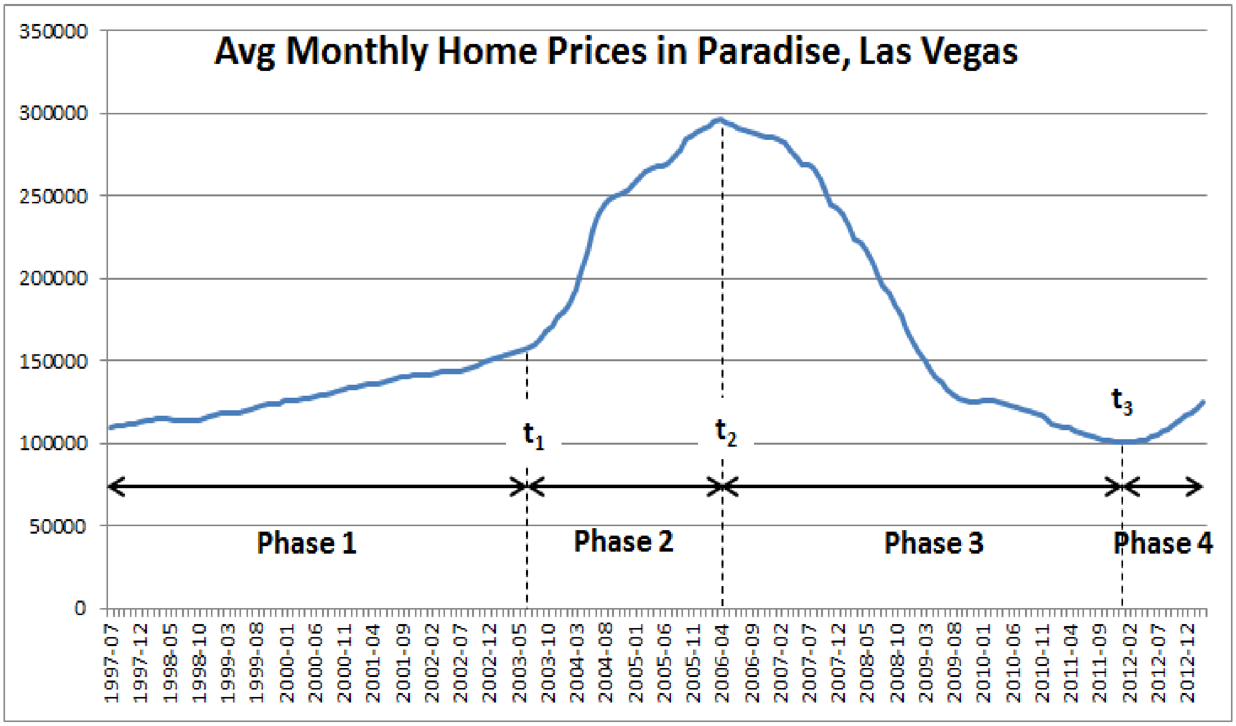

Figure 7: Average monthly house prices in Paradise, Las Vegas

We plotted the data in Figure 7 and found the following: in phase 1 from 1997 through 2003, there is a normal increase in house prices, and then quite suddenly in 2003 there is a phenomenal increase during which many jumped into the housing market. Then, starting in 2006 for 6 years house prices had a significant drop. Finally, in early 2012 house prices start rising modestly, more or less at the same level as back in 1997.

3.2 LITERATURE REVIEW

3.2.1 Previous Related Work On Home Mortgage Default

Ciurlia and Gheno [6] present a two-factor model where the real estate asset value and the spot rate dynamics are jointly modeled in order to take into account the real estate market sensitivity to the interest rate term structure. The pricing problem for both European and American options is then analyzed and a discrete-time bi-dimensional binomial lattice framework is adopted. However, in this paper, only the price of the option is calculated numerically. No analytic approximation is derived and no foreclosure probability is derived. Kuo [15] criticizes [6] on the use of geometric Brownian motion. The author proposes a polynomial approximation method to value the mortgage default option. The author uses actual transaction data to estimate a more realistic price process and applies the empirically estimated house price model to value the default option rather than assuming the house price to be a random walk process. The author sets up a house price model with three return components: i.e., the AR(1) market return, the AR(1) persistent idiosyncratic error, and the time-independent transaction error. However, the data set being used is too old which is from a paper by Case and Shiller [8]. Also, only the price of mortgage default option has been calculated. No default probability is derived.

Gelain and Lansing [9] investigate the behavior of the equilibrium price-rent ratio for housing in a standard asset pricing model and compare the model predictions to survey evidence on the return expectations of real-world housing investors. The authors show that if agents in the model employ simple moving-average forecast rules, they tend to expect higher future returns when house prices are high relative to fundamentals. This paper only discusses the mean-reversion property of house price itself. However, it does not discuss the options on houses, including prices and default probabilities.

Krainer, LeRoy, and O [16] develop an equilibrium valuation model that incorporates optimal default to show how mortgage yields and lender recovery rates on defaulted mortgages depend on initial loan-to-value (LTV) ratios. The analysis treats both the frictionless case and the case in which borrowers and lenders incur deadweight costs upon default. As the model (or common sense) predicts, high-LTV mortgages are more prone to default than low-LTV mortgages. Further, the dependence of mortgage pricing on LTV conforms to the prediction of the model. This paper only analyzes the effect of LTV on default and mortgage pricing. However, it does not discuss the options on houses.

Willen, Foote, Gerardi, and Goette [21] from the Federal Reserve Bank of Boston discuss possible reasons to cause the foreclosure crisis and what might be done to stop it. They focus on two key decisions: the borrower's choice to default on the mortgage and the lender's choice on whether to renegotiate or "modify" the loan. However, very little work has been done in the above paper or elsewhere on the quantitative modeling of these strategies.

Ambrose and Buttimer [2] propose a discrete version of the Adjustable Balance Mortgage (ABM), which allows the homeowner to pay a reduced mortgage payment if the house price falls below the remaining value of the mortgage. However, the analysis is in discrete time and is difficult to compare with the alternative approaches to be described here that are carried out in the continuous time setting.

Our purpose is to use the simplest model, which is able to capture the phenomenon and quantitatively study some of the strategies discussed in the two previous papers. Both analytic approximations of the option price and default (foreclosure) probability are derived. Moreover, the regime switching case is handled.

3.2.2 Contribution

In this second part of the dissertation, we explore two different ways to avoid mortgage defaults when housing prices decline. First, we propose a new mortgage insurance contract (American put option with the house as the underlying asset). Second, we price the adjustable balance mortgage in continuous time, in order to facilitate the comparison of the two approaches and to make possible the calculation of probabilities of foreclosures.

The put problem has two essential differences from the more standard equity case. First, the rate of the randomly priced stock (the underlier) is replaced by the house which the owner can "sell" (exchange) at any time for the remaining value of the mortgage, $M(t)$. Moreover, the house price has a regime switch from falling to rising values. The second essential difference is that the strike price of the option, $M(t)$, is now a function of t.

3.3 MATHEMATICAL MODEL OF HOME FORECLOSURE

In this section we will develop a simplified mathematical model of the most common type of foreclosure. Specifically, the homeowner stops paying the mortgage and vacates the house the instant that the house price falls below the remaining value of the mortgage (i.e. is underwater). Of course, there are many extenuating circumstances that preclude the homeowners from foreclosing at this precise instant - e.g., employment, children's schooling, availability of alternative housing, etc. On the other hand, this ambiguity of the precise moment of default arises, and is common, in many other situations: defaults of corporate bonds, credit instruments and (as pointed out in the previous chapters) financial institutions. Another simplification in our study is that the analysis will be carried out in the continuous time setting rather than the monthly payment structure common to most mortgages. This will make the analysis simpler and the results more transparent especially when we study the early exercise boundaries and the first crossing problems that are key to our work.

3.3.1 Mathematical Models Of House Prices and Mortgage Values

We assume the homeowner has obtained a fixed rate mortgage (FRM) with rate c, and with maturity T^* (say T^* is 30 years). This fixed rate mortgage satisfies the following differential equation:

$$dM(t) = cM(t)dt - mdt, \quad M(T^*) = 0 \tag{3.3.1}$$

where m is the continuous repayment rate. It follows that

$$d\left(e^{-ct}M(t)\right) = -me^{-ct}dt \tag{3.3.2}$$

and integrating both sides from t to T^*, we have

$$e^{-cT^*}M(T^*) - e^{-ct}M(t) = -m\int_t^{T^*} e^{-ct}dt \tag{3.3.3}$$

Since $M(T^*) = 0$, we have

$$-e^{-ct}M(t) = \frac{m}{c}(e^{-cT^*} - e^{-ct}) \tag{3.3.4}$$

Thus

$$M(t) = \frac{m}{c}\left(1 - e^{-c(T^*-t)}\right) \tag{3.3.5}$$

with

$$M(0) = \frac{m}{c}\left(1 - e^{-cT^*}\right) \tag{3.3.6}$$

or

$$\frac{m}{c} = \frac{M(0)}{1 - e^{-cT^*}} \tag{3.3.7}$$

In summary,

$$M(t) = \frac{m}{c}\left(1 - e^{-c(T^*-t)}\right) = M(0)\left(\frac{1 - e^{-c(T^*-t)}}{1 - e^{-cT^*}}\right) \tag{3.3.8}$$

Individual house prices are modeled by a Geometric Brownian Motion that satisfies a stochastic differential equation of the form

$$dS = (r - \delta(t))Sdt + \sigma(t)SdW(t) \tag{3.3.9}$$

where $W(t)$ is a Brownian motion and the drift and volatility are given by step functions

$$\delta(t) = \delta_1\chi_{[0,T_s]} + \delta_2\chi_{[T_s,T]}, \qquad \delta_2 < r < \delta_1 \tag{3.3.10}$$

$$\sigma(t) = \sigma_1\chi_{[0,T_s]} + \sigma_2\chi_{[T_s,T]} \tag{3.3.11}$$

where T_s is the time the regime switches from falling prices to rising prices. Here χ_A is the indicator function of the set {A} defined as $\chi_A(x) = \begin{cases} 1 & if x \in A \\ 0 & if x \notin A \end{cases}$.

We assume that individual house prices follow the average in the region (i.e., we will fit the parameters in Equation 3.3.9 to the data in Table 8). To this end, we find, using standard techniques [20], that $\delta_1 = 0.12$, $\delta_2 = 0.03$, $\sigma_1 = 0.045767598$, $\sigma_2 = 0.031712223$ provide a reasonable fit to the Zillow data and we take $r = 0.05$ and $c = 0.06$ from historical data. One rather surprising observation is that the volatility is quite low in both the falling and

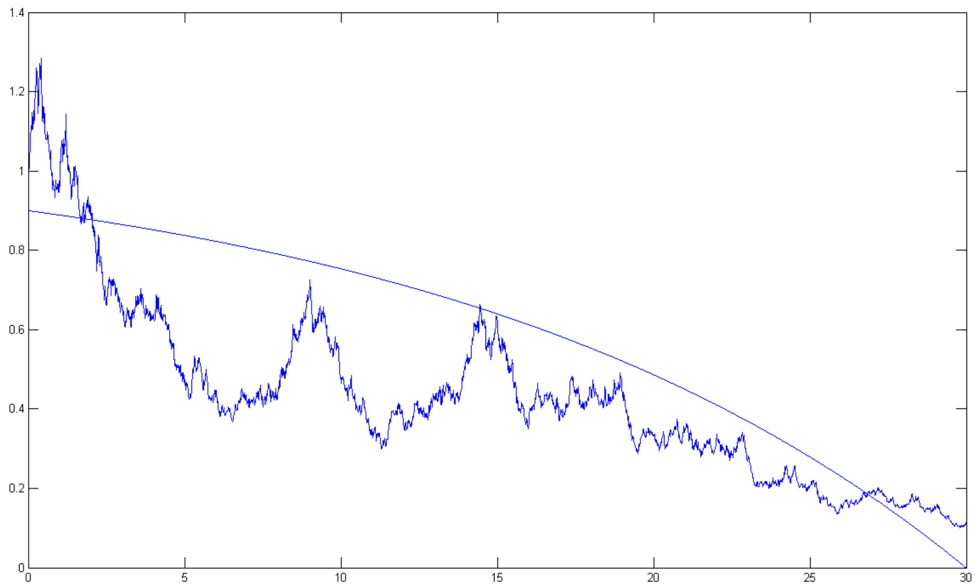

Figure 8: Common type of default

rising house price regimes.

Figure 8 shows a single sample path, $S(t)$, of Equation 3.3.9 (the jagged curve) with initial house price normalized to 1 along with the remaining value of the mortgage, $M(t)$, (the smooth curve with the initial value of the mortgage, $M(0) = 0.9$; i.e. 90% of the house price). The mathematical definition of the common type of default (foreclosure) described above is

$$\tau = \inf \{t > 0; S(t) \leq M(t)\} \tag{3.3.12}$$

Both the homeowner, and the mortgage company, are interested in the default probability (probability of foreclosure) defined mathematically by $P(\tau \leq t)$.

3.3.2 Probability Of Foreclosure

In this section we calculate the default probability defined above for the model (Equation 3.3.8 and 3.3.9). We accomplish this by rephrasing the problem as a first crossing problem for Brownian motions [20].

In phase 1, $S(t) = S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t+\sigma_1 W(t)}$ so that $S(t) \leq M(t)$ is equivalent to

$$S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t+\sigma_1 W(t)} \leq \frac{M(0)}{1-e^{-cT^*}}\left(1-e^{-c(T^*-t)}\right) \tag{3.3.13}$$

$$\sigma_1 W(t) \leq \ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - (r-\delta_1-\frac{\sigma_1^2}{2})t \tag{3.3.14}$$

Explicit solutions to first crossing problems for Brownian motions are only known for lines; i.e., $P(W(t) \leq A + Bt)$ [19]. To fit this requirement we approximate the term $\ln\left(1-e^{-c(T^*-t)}\right) \approx -e^{-c(T^*-t)} \approx -e^{-cT^*}e^{ct} \approx -e^{-cT^*}(1+ct)$. This results in:

$$W(t) \leq \frac{1}{\sigma_1}\left\{\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} - e^{-cT^*} - \left(r-\delta_1-\frac{\sigma_1^2}{2}+ce^{-cT^*}\right)t\right\} \tag{3.3.15}$$

Suppose $X(t) = \mu t + \sigma W(t)$, where $W(t)$ is a Brownian Motion. We define the first passage time that the process $X(t)$ crosses the line $y = b$ from above as

$$\begin{aligned}\tau &= \inf\left\{t>0, X(t)\leq b\right\} = \inf\left\{t>0, \mu t+\sigma W(t)\leq b\right\}\\ &= \inf\left\{t>0, W(t)\leq \frac{b}{\sigma}-\frac{\mu}{\sigma}t\right\} = \inf\left\{t>0, W(t)\leq B+At\right\}\end{aligned} \tag{3.3.16}$$

where $A = -\frac{\mu}{\sigma}$ and $B = \frac{b}{\sigma}$.

The survival probability can be expressed in terms of its pdf [19].

$$P(\tau > t) = \int_{B+At}^{+\infty} u(x,t)dx \tag{3.3.17}$$

where $u(x,t) = \frac{1}{\sqrt{2\pi t}}\left[e^{-\frac{x^2}{2t}} - e^{-2AB}e^{-\frac{(2B-x)^2}{2t}}\right]$. Thus

$$\begin{aligned}P(\tau > t) &= \frac{1}{\sqrt{2\pi t}}\int_{B+At}^{+\infty} e^{-\frac{x^2}{2t}}dx - \frac{1}{\sqrt{2\pi t}}\int_{B+At}^{+\infty} e^{-2AB}e^{-\frac{(2B-x)^2}{2t}}dx\\ &= \frac{1}{\sqrt{2\pi}}\int_{\frac{B}{\sqrt{t}}+A\sqrt{t}}^{+\infty} e^{-\frac{y^2}{2}}dy - \frac{1}{\sqrt{2\pi}}e^{-2AB}\int_{-\frac{B}{\sqrt{t}}+A\sqrt{t}}^{+\infty} e^{-\frac{y^2}{2}}dy\\ &= N\left(-\frac{B}{\sqrt{t}}-A\sqrt{t}\right) - e^{-2AB}N\left(\frac{B}{\sqrt{t}}-A\sqrt{t}\right)\\ &= N\left(-\frac{b}{\sigma\sqrt{t}}+\frac{\mu}{\sigma}\sqrt{t}\right) - e^{\frac{2b\mu}{\sigma^2}}N\left(\frac{b}{\sigma\sqrt{t}}+\frac{\mu}{\sigma}\sqrt{t}\right)\end{aligned} \tag{3.3.18}$$

where N is the normalized Gaussian CDF. Therefore, in phase 1 (i.e. when $0 \leq t \leq T_s$), we have

$$P(\tau > t) = N\left(-\frac{b_1}{\sigma_1\sqrt{t}} + \frac{\mu_1}{\sigma_1}\sqrt{t}\right) - e^{\frac{2b_1\mu_1}{\sigma_1^2}} N\left(\frac{b_1}{\sigma_1\sqrt{t}} + \frac{\mu_1}{\sigma_1}\sqrt{t}\right) \tag{3.3.19}$$

where $b_1 = \ln \frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} - e^{-cT^*}$ and $\mu_1 = r - \delta_1 - \frac{\sigma_1^2}{2} + ce^{-cT^*}$

In phase 2 (i.e. when $T_s \leq t \leq T$) we have $S(t) = S(T_s)e^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)}$

so that

$$S(t) = S(T_s)e^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \leq M(t) = \frac{M(0)}{1-e^{-cT^*}}\left(1 - e^{-c(T^*-t)}\right) \tag{3.3.20}$$

is equivalent to

$$\left(r - \delta_2 - \frac{\sigma_2^2}{2}\right)(t - T_s) + \sigma_2 W(t - T_s) \leq \ln \frac{M(0)}{S(T_s)\left(1 - e^{-cT^*}\right)} + \ln\left(1 - e^{-c(T^*-t)}\right) \tag{3.3.21}$$

Again, approximating $\ln\left(1 - e^{-c(T^*-t)}\right) \approx -e^{-c(T^*-t)} \approx -e^{-cT^*}e^{ct} \approx -e^{-cT^*}(1 + ct)$, this inequality is approximated by

$$W(t-T_s) \leq \frac{1}{\sigma_2}\left\{\ln \frac{M(0)}{S(T_s)\left(1 - e^{-cT^*}\right)} - e^{-cT^*} - cT_s e^{-cT^*} - (r - \delta_2 - \frac{\sigma_2^2}{2} + ce^{-cT^*})(t - T_s)\right\} \tag{3.3.22}$$

Therefore, in phase 2 with $T_s \leq t \leq T$, using Equation 3.3.19 in the first term below and letting the random variable $S(T_s) = y$, we have

$$\begin{aligned} &P(\tau > t) \\ &= P(\tau > T_s)\left(\int_{M(T_s)}^{+\infty} f(y)dy\right)^{-1} \\ &\int_{M(T_s)}^{+\infty}\left\{N\left(-\frac{b_2}{\sigma_2\sqrt{t-T_s}} + \frac{\mu_2}{\sigma_2}\sqrt{t-T_s}\right) - e^{\frac{2b_2\mu_2}{\sigma_2^2}} N\left(\frac{b_2}{\sigma_2\sqrt{t-T_s}} + \frac{\mu_2}{\sigma_2}\sqrt{t-T_s}\right)\right\} f(y)dy \end{aligned} \tag{3.3.23}$$

where $f(y) = \frac{1}{\sigma_1 y\sqrt{2\pi T_s}} e^{-\frac{\left(\ln\frac{y}{S(0)} - (r-\delta_1-\frac{\sigma_1^2}{2})T_s\right)^2}{2\sigma_1^2 T_s}}$ is the transition probability density function $p(S(0), 0; y, T_s)$ for Geometric Brownian Motions [20], and the constants are given by

$b_2 = \ln \frac{M(0)}{y\left(1-e^{-cT^*}\right)} - e^{-cT^*} - cT_s e^{-cT^*}$, $\mu_2 = r - \delta_2 - \frac{\sigma_2^2}{2} + ce^{-cT^*}$, and

$M(T_s) = \frac{M(0)}{1-e^{-cT^*}}\left(1 - e^{-c(T^*-T_s)}\right)$.

The normalizing factor $\left(\int_{M(T_s)}^{+\infty} f(y)dy\right)^{-1}$ is included to give a probability of 1 to the sample paths that survived to T_s.

Using the parameter values from the Zillow data in Equation 3.3.9 and 3.3.8, we calculate numerically the survival probability $P(\tau > t) = 1 - P(\tau \leq t)$ using Equation 3.3.19 and 3.3.23, and plot the results in Figure 9.

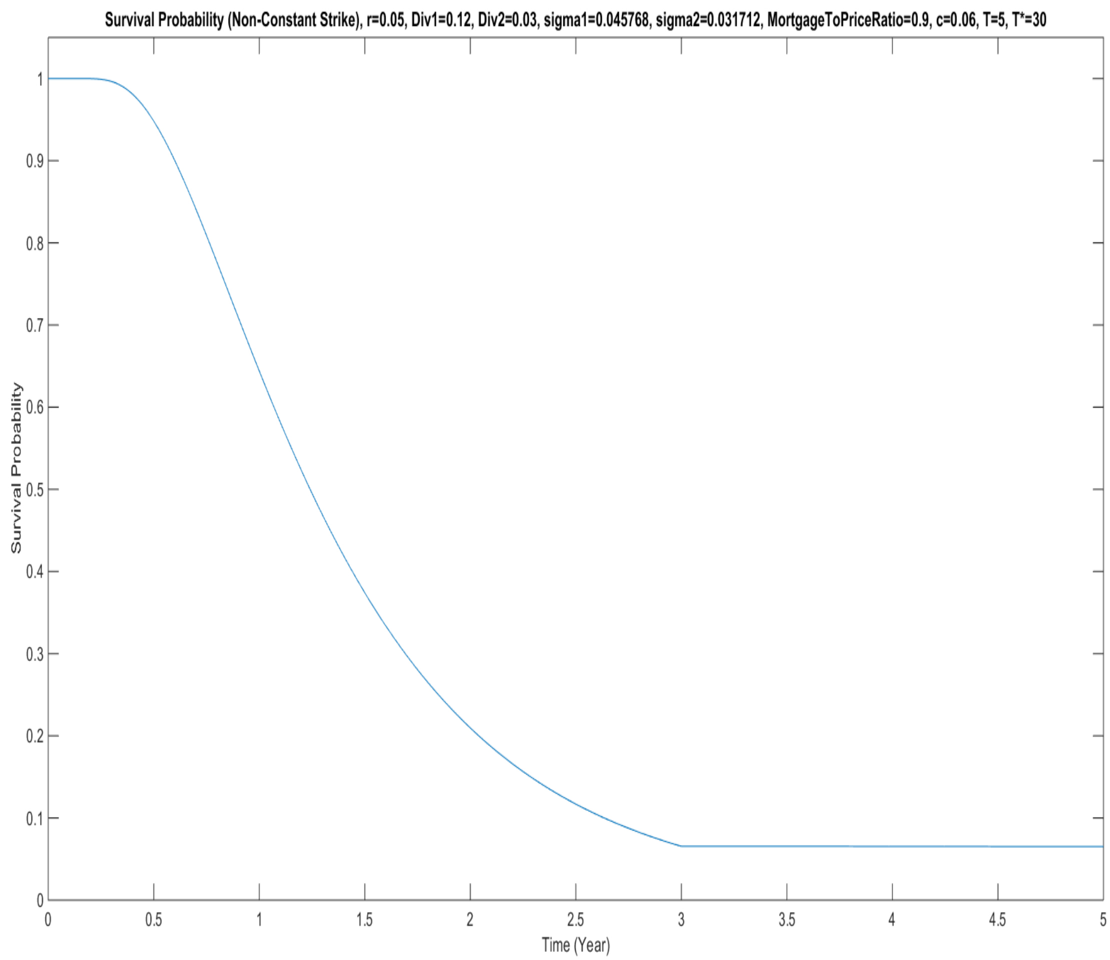

Figure 9: Survival probability for common foreclosure model. (Strike = $M(t)$, $r = 0.05,\ \delta_1 = 0.12,\ \delta_2 = 0.03,\ \sigma_1 = 0.045768,\ \sigma_2 = 0.031712$, Mortgage To Price Ratio $= 0.9,\ c = 0.06,\ T = 5,\ T_s = 3,\ T^* = 30$)

It is clear from Figure 9 that the probability of foreclosure begins to rapidly increase after the first half-year and approaches 90% before house prices begin to rise at $T_s = 3$. After T_s there is a very modest decrease in the survival probability, as anticipated, since house prices have now begun to rise. However by this time a significant proportion of homeowners would have foreclosed on their mortgage if they had followed the simple criterion that their house value had fallen below the remaining value of their mortgage. As mentioned, this would have a significant negative impact on the homeowner's future credit rating and the mortgage company would be forced to hold a non-performing mortgage as well as to acquire a devalued house.

In the next section we will introduce two financial strategies to avoid, or reduce the frequency of foreclosures of this type.

3.4 PUT OPTION ON THE MORTGAGE

The first strategy to reduce foreclosure frequency is to allow the homeowner to buy a put option on his home allowing him (her) to sell (exchange) it for the remaining value of the mortgage without future credit consequences. The amount the homeowner would pay upfront for being able to legally walk away from his house and mortgage would be determined by the mortgage company using risk-neutral pricing so that the financial risk of this type of foreclosure is equally shared by the two parties.

3.4.1 The American Put Problem

American style put options are well known in the equity options literature. This instrument gives the holder the right to sell any time in the future a stock for some constant value which has been decided today. Our put has two essential differences from the equity case. First, the rate of the randomly priced stock (the underlier) is replaced by the house which the owner can "sell" (exchange) at any time for the remaining value of the mortgage, $M(t)$. Moreover, Equation 3.3.9 modeling this house price has a regime switch from falling to rising values.

The second essential difference is that the strike price of the option, $M(t)$, is now a function of t.

All options will be priced using the Black-Scholes, risk-neutral formalism. We begin by stating the mathematical problem for P, the price of the American put option. At maturity of the put, say $T = 5$ years in the future, the intrinsic payoff is the difference between the remaining mortgage value and the house price. But an American style option allows a person to exercise at any time between the purchase of the option and its maturity. Thus we must incorporate the optimal early exercise boundary $B(t)$, namely, the value of the house price dilutes below which it is optimal for the person who bought this put option to foreclose on the house by exchanging it for the mortgage. Specifically the price, P, at time 0, for the American put option to exchange the house for the remaining value of the mortgage satisfies the Black-Scholes PDE [20].

$$P_t + \frac{1}{2}\sigma^2(t)S^2 P_{SS} + (r - \delta(t))SP_S - rP = 0, \quad S > B(t), \quad 0 < t < T \tag{3.4.1}$$

$$P(B(t), t) = M(t) - B(t) \tag{3.4.2}$$

$$P_S(B(t), t) = -1 \tag{3.4.3}$$

$$P(S, T) = (M(T) - S)^+ \tag{3.4.4}$$

$$B(T) = min\left(M(T), \frac{r}{\delta_2}M(T)\right) \tag{3.4.5}$$

Free boundary problems such as Equation 3.4.1 to Equation 3.4.5, even when $\delta(t)$, $\sigma(t)$, and $M(t)$ are constants, do not have explicit solutions - they must be solved numerically. Following Carr, Jarrow, and Myneni [7] we obtain an integral representation for the price, P_{Am}, in terms of the unknown early exercise boundary, $B(t)$. Essentially, the method decomposes the value of an American put option into the corresponding European put price and the early exercise premium. In the non-regime switching case (δ, σ are constant, for example as in phase 1), the Carr integral equation is [7]:

$$
\begin{aligned}
&P_{Am}(S,t) = P_{Eu}(S,t) \\
&+ \int_t^T \left(me^{-c(T^*-u)} + rM(u)\right) e^{-r(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{S}\right) - (r-\delta-\frac{\sigma^2}{2})(u-t)\right)\right) du \\
&- \delta S \int_t^T e^{-\delta(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{S}\right) - (r-\delta+\frac{\sigma^2}{2})(u-t)\right)\right) du
\end{aligned}
\tag{3.4.6}
$$

where $N(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$ is the standard normal CDF. Here the variable strike appears in the first integral.

The standard approach is to first solve for the early exercise boundary. This is accomplished by evaluating Equation 3.4.6 on the boundary, i.e., by replacing S by $B(t)$ and using Equation 3.4.2 to obtain a nonlinear integral equation for $B(t)$. This generalizes Carr's integral equation when the strike is not constant, and as we shall see, it can be extended to the case of interest here when $\delta(t)$ and $\sigma(t)$ are functions of t. We then solve this integral equation numerically to find the early exercise boundary $B(t)$. Finally, $B(t)$ can be inserted into Equation 3.4.6 to obtain the values of P_{Am}.

With this instrument, the mortgage holder will not necessarily default the minute his house price goes below the mortgage value (as in the previous section), but rather will follow the optimal strategy to see if $S(t)$ drops below $B(t)$.

On $S = B(t)$ (approaching from the continuation region where the PDE holds and using Equation 3.4.4), Equation 3.4.6 becomes

$$
\begin{aligned}
&P_{Am}(B(t),t) = M(t) - B(t) = P_{Eu}(B(t),t;M(T),T) \\
&+ \int_t^T \left(me^{-c(T^*-u)} + rM(u)\right) e^{-r(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r-\delta-\frac{\sigma^2}{2})(u-t)\right)\right) du \\
&- \delta B(t) \int_t^T e^{-\delta(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r-\delta+\frac{\sigma^2}{2})(u-t)\right)\right) du
\end{aligned}
\tag{3.4.7}
$$

We will solve this integral equation numerically (and ultimately incorporate the regime switching as well). But in order to be confident that our numerical scheme gives correct results, we will first develop the code in the constant strike case, $K = M(T)$, and compare our results with the existing results of Chen, Cheng, and Chadam [5].

For constant strike, Equation 3.4.7 reduces to

$$\begin{aligned} K - B(t) =& P_{Eu}(B(t), t; M(T), T) \\ &+ K \int_t^T r e^{-r(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r - \delta - \frac{\sigma^2}{2})(u-t)\right)\right) du \\ &- \delta B(t) \int_t^T e^{-\delta(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r - \delta + \frac{\sigma^2}{2})(u-t)\right)\right) du \end{aligned} \tag{3.4.8}$$

We will study Equation 3.4.8 for two sets of parameters: for $r > \delta$ (rising house prices), and for $r < \delta$ (falling house prices) since for American put options, these two cases are different at expiry [14]:

$$B(T) = \begin{cases} K & \text{if } \delta \leq r \\ \frac{r}{\delta}K & \text{if } \delta \geq r \end{cases} \tag{3.4.9}$$

An outline of the numerical scheme for $B(t)$ follows. On $[0, T]$, solve the integral equation for $B(t)$ numerically, i.e. take $0 = t_0 < t_1 < ... < t_n = T$, with $\Delta t = \frac{T}{n}$. Denote by B_k the resulting numerical approximation for $B(t_k)$. Start with $B_n = B(T) = \begin{cases} K & \text{if } \delta \leq r \\ \frac{r}{\delta}K & \text{if } \delta \geq r \end{cases}$

At the first time step $t_{n-1} = T - \Delta t$, we proceed as follows. Take $B^0_{n-1} = B_n = B(T)$ as the initial guess. Then numerically solve for B_{n-1} in the non-linear equation. We use right Riemann sums to approximate the two integrals. We have:

$$\begin{aligned}
B_{n-1} &= K - P_{Eu}(B_{n-1}, t_{n-1}; K, T) \\
&- K \int_{t_{n-1}}^{t_n} re^{-r(u-t_{n-1})} N\left(\frac{1}{\sigma\sqrt{u-t_{n-1}}}\left(\ln\left(\frac{B(u)}{B_{n-1}^0}\right) - (r-\delta-\frac{\sigma^2}{2})(u-t_{n-1})\right)\right) du \\
&+ B_{n-1}^0 \int_{t_{n-1}}^{t_n} \delta e^{-\delta(u-t_{n-1})} N\left(\frac{1}{\sigma\sqrt{u-t_{n-1}}}\left(\ln\left(\frac{B(u)}{B_{n-1}^0}\right) - (r-\delta+\frac{\sigma^2}{2})(u-t_{n-1})\right)\right) du \\
&= K - P_{Eu}(B_{n-1}, t_{n-1}; K, T) \\
&- Kre^{-r\Delta t} N\left(\frac{1}{\sigma\sqrt{\Delta t}}\left(\ln\left(\frac{B_n}{B_{n-1}^0}\right) - (r-\delta-\frac{\sigma^2}{2})\Delta t\right)\right) \Delta t \\
&+ B_{n-1}^0 \delta e^{-\delta\Delta t} N\left(\frac{1}{\sigma\sqrt{\Delta t}}\left(\ln\left(\frac{B_n}{B_{n-1}^0}\right) - (r-\delta+\frac{\sigma^2}{2})\Delta t\right)\right) \Delta t
\end{aligned} \tag{3.4.10}$$

At the second time step $t_{n-2} = T - 2\Delta t$, we use the initial guess $B_{n-2}^0 = B_{n-1}$, so we have:

$$\begin{aligned}
B_{n-2} &= K - P_{Eu}(B_{n-2}, t_{n-2}; K, T) \\
&- K \int_{t_{n-2}}^{t_{n-1}} re^{-r(u-t_{n-2})} N\left(\frac{1}{\sigma\sqrt{u-t_{n-2}}}\left(\ln\left(\frac{B(u)}{B_{n-2}^0}\right) - (r-\delta-\frac{\sigma^2}{2})(u-t_{n-2})\right)\right) du \\
&- K \int_{t_{n-1}}^{t_n} re^{-r(u-t_{n-2})} N\left(\frac{1}{\sigma\sqrt{u-t_{n-2}}}\left(\ln\left(\frac{B(u)}{B_{n-2}^0}\right) - (r-\delta-\frac{\sigma^2}{2})(u-t_{n-2})\right)\right) du \\
&+ B_{n-2}^0 \int_{t_{n-2}}^{t_{n-1}} \delta e^{-\delta(u-t_{n-2})} N\left(\frac{1}{\sigma\sqrt{u-t_{n-2}}}\left(\ln\left(\frac{B(u)}{B_{n-2}^0}\right) - (r-\delta+\frac{\sigma^2}{2})(u-t_{n-2})\right)\right) du \\
&+ B_{n-2}^0 \int_{t_{n-1}}^{t_n} \delta e^{-\delta(u-t_{n-2})} N\left(\frac{1}{\sigma\sqrt{u-t_{n-2}}}\left(\ln\left(\frac{B(u)}{B_{n-2}^0}\right) - (r-\delta+\frac{\sigma^2}{2})(u-t_{n-2})\right)\right) du
\end{aligned} \tag{3.4.11}$$

i.e.

$$\begin{aligned}
B_{n-2} =& K - P_{Eu}(B_{n-2}, t_{n-2}; K, T) \\
&- Kre^{-r\Delta t} N\left(\frac{1}{\sigma\sqrt{\Delta t}}\left(\ln\left(\frac{B_{n-1}}{B_{n-2}^0}\right) - (r-\delta-\frac{\sigma^2}{2})\Delta t\right)\right) \Delta t \\
&- Kre^{-r2\Delta t} N\left(\frac{1}{\sigma\sqrt{2\Delta t}}\left(\ln\left(\frac{B_n}{B_{n-2}^0}\right) - (r-\delta-\frac{\sigma^2}{2})2\Delta t\right)\right) \Delta t \\
&+ B_{n-2}^0 \delta e^{-\delta\Delta t} N\left(\frac{1}{\sigma\sqrt{\Delta t}}\left(\ln\left(\frac{B_{n-1}}{B_{n-2}^0}\right) - (r-\delta+\frac{\sigma^2}{2})\Delta t\right)\right) \Delta t \\
&+ B_{n-2}^0 \delta e^{-\delta 2\Delta t} N\left(\frac{1}{\sigma\sqrt{2\Delta t}}\left(\ln\left(\frac{B_n}{B_{n-2}^0}\right) - (r-\delta+\frac{\sigma^2}{2})2\Delta t\right)\right) \Delta t
\end{aligned} \tag{3.4.12}$$

In general, at the j^{th} time step $t_{n-j} = T - j\Delta t$, we use the initial guess $B^0_{n-j} = B_{n-j+1}$ and solve the equation:

$$\begin{aligned}B_{n-j} =& K - P_{Eu}(B_{n-j}, t_{n-j}; K, T) \\ & - \sum_{s=1}^{j} Kre^{-rs\Delta t} N\left(\frac{1}{\sigma\sqrt{s\Delta t}}\left(\ln\left(\frac{B_{n-j+s}}{B^0_{n-j}}\right) - (r - \delta - \frac{\sigma^2}{2})s\Delta t\right)\right)\Delta t \\ & + \sum_{m=1}^{j} B^0_{n-j}\delta e^{-\delta m\Delta t} N\left(\frac{1}{\sigma\sqrt{m\Delta t}}\left(\ln\left(\frac{B_{n-j+m}}{B^0_{n-j}}\right) - (r - \delta + \frac{\sigma^2}{2})m\Delta t\right)\right)\Delta t\end{aligned} \tag{3.4.13}$$

Following this procedure, we obtain the solution B_{n-j} at each time step t_{n-j}. With sufficiently small time steps, we can accurately approximate the optimal early exercise boundary for the American put option.

In the regime switching case (two phases), during the second phase, we can calculate the integral completely as above, but before the regime switching time T_s, the integrals must be broken into two pieces each, and both values of the dividend rate, δ_1 and δ_2, will appear. Specifically, in the regime switching case Equation 3.4.6 becomes

$$\begin{aligned}& P_{Am}(S,t) = P_{Eu}(S,t) \\ & + \int_{T_s}^{T} \left(me^{-c(T^*-u)} + rM(u)\right) e^{-r(u-t)} \\ & N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{S}\right) - (r - \frac{\sigma^2}{2})(u-t) + \delta_1(T_s - t) + \delta_2(u - T_s)\right)\right) du \\ & + \int_{t}^{T_s} \left(me^{-c(T^*-u)} + rM(u)\right) e^{-r(u-t)} \\ & N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{S}\right) - (r - \frac{\sigma^2}{2})(u-t) + \delta_1(u-t)\right)\right) du \\ & - \delta_2 S \int_{T_s}^{T} e^{-(\delta_1(T_s-t)+\delta_2(u-T_s))} \\ & N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{S}\right) - (r + \frac{\sigma^2}{2})(u-t) + \delta_1(T_s - t) + \delta_2(u - T_s)\right)\right) du \\ & - \delta_1 S \int_{t}^{T_s} e^{-\delta_1(u-t)} \\ & N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{S}\right) - (r + \frac{\sigma^2}{2})(u-t) + \delta_1(u-t)\right)\right) du\end{aligned} \tag{3.4.14}$$

On $S = B(t)$, with constant strike K, we have:

$$
\begin{aligned}
K - B(t) &= P_{Eu}(B(t), t; K, T) \\
&+ K \int_{T_s}^{T} r e^{-r(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r - \frac{\sigma^2}{2})(u-t) + \delta_1(T_s - t) + \delta_2(u - T_s)\right)\right) \\
&\quad du \\
&+ K \int_{t}^{T_s} r e^{-r(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r - \frac{\sigma^2}{2})(u-t) + \delta_1(u - t)\right)\right) du \\
&- \delta_2 B(t) \int_{T_s}^{T} e^{-(\delta_1(T_s - t) + \delta_2(u - T_s))} \\
&\qquad N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r + \frac{\sigma^2}{2})(u-t) + \delta_1(T_s - t) + \delta_2(u - T_s)\right)\right) du \\
&- \delta_1 B(t) \int_{t}^{T_s} e^{-\delta_1(u-t)} N\left(\frac{1}{\sigma\sqrt{u-t}}\left(\ln\left(\frac{B(u)}{B(t)}\right) - (r + \frac{\sigma^2}{2})(u-t) + \delta_1(u - t)\right)\right) du
\end{aligned}
\tag{3.4.15}
$$

This integral equation is solved numerically using a scheme similar to 3.4.13.

3.4.2 Numerical Validation Of The Scheme

The non-regime switching case with constant strike K, and with constant $r > \delta$, is the well studied standard case [7]. We present the numerical results with the following parameters: $S(0) = 1$ (We normalize the house price to 1 and take the constant strike $K = 0.8376$ to be the value of $M(5)$ for a 6%, 30-year fixed rate mortgage, i.e., $c = 0.06$, $T^* = 30$ in 3.3.8 with the mortgage to house price ratio $M(0) = 0.9$); i.e., $K = 0.8376$; and the put option maturity in years is $T = 5$; $r = 0.05$; $\delta = 0.03$; $\sigma = 0.2$; $dt = \frac{1}{200}$ (in year). Figure 10 shows the perfect match of the scheme 3.4.13 with previous calculations using other methods [5].

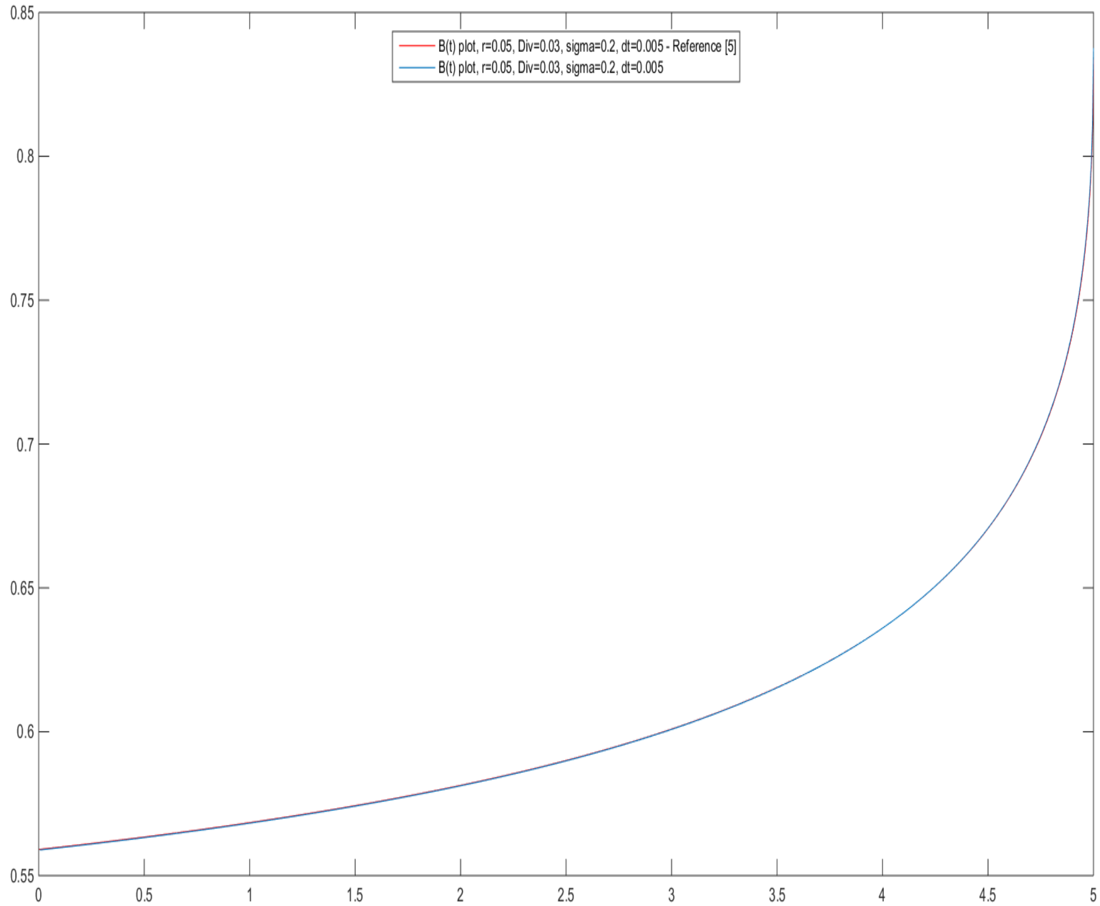

Figure 10: Comparison (with [5]) of the early exercise boundary for constant strike. $(K = 0.8376,\ r = 0.05,\ \delta = 0.03,\ \sigma = 0.2,\ dt = 0.005,\ T = 5)$

We also carry out the calculations for falling house prices with $r < \delta$. This is the case where the boundary does not start at the strike K, but rather below the strike at $\frac{r}{\delta}K$. Specifically, $S(0) = 1$; $K = 0.8376$; $T = 5$; $r = 0.05$; $\delta = 0.12$; $\sigma = 0.2$; $dt = \frac{1}{200}$ (in year). In this case, at expiry the early exercise boundary begins at $\frac{r}{\delta}K = 0.349$ rather than at $K = 0.8376$. Once again, our results match well. See Figure 11 for details.

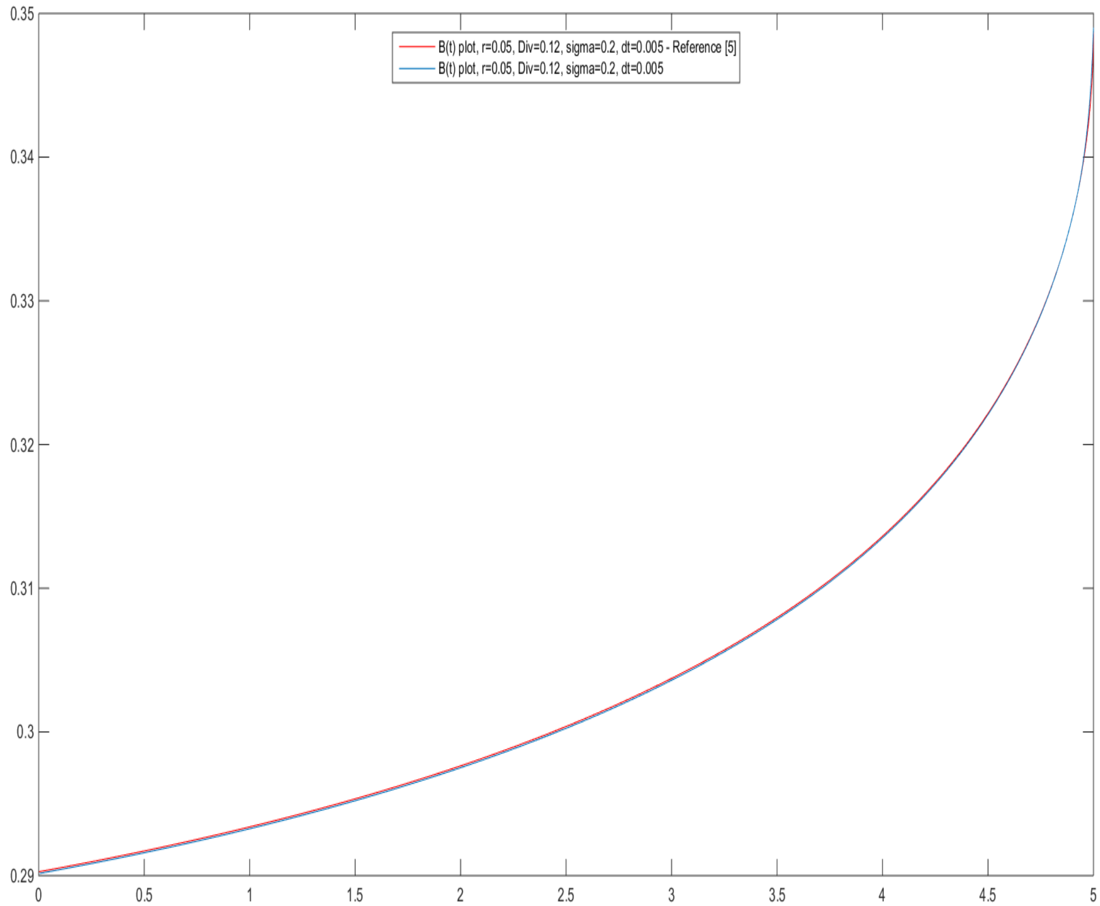

Figure 11: Comparison (with [5]) of the early exercise boundary for constant strike. ($K = 0.8376,\ \frac{r}{\delta}K = 0.349,\ r = 0.05,\ \delta = 0.12,\ \sigma = 0.2,\ dt = 0.005,\ T = 5$)

Finally, in the regime switching case, with constant strike, we use the parameters: $S(0) = 1$; $K = 0.8376$; $T = 5$; $T_s = 3$ (Regime switching time, in years); $r = 0.05$; $\delta_1 = 0.12$; $\delta_2 = 0.03$; $\sigma = 0.2$; $dt = \frac{1}{200}$ (in year). From $T_s = 3$ to expiry, $T = 5$, the house prices are rising because the dividend rate is 3% and the interest rate is 5%. In the earlier period, $0 \leq t \leq T_s = 3$, (phase 1), when the house prices are falling, the dividend rate is 12%. From Figure 12, one notices a jump in the early exercise boundary, $B(t)$, at the regime switching time due to the dividend rate change. Following the analysis of Jiang [14] the starting point of the boundary at T_s is $min\left(\frac{r}{\delta_1}K, B(T_s+)\right)$.

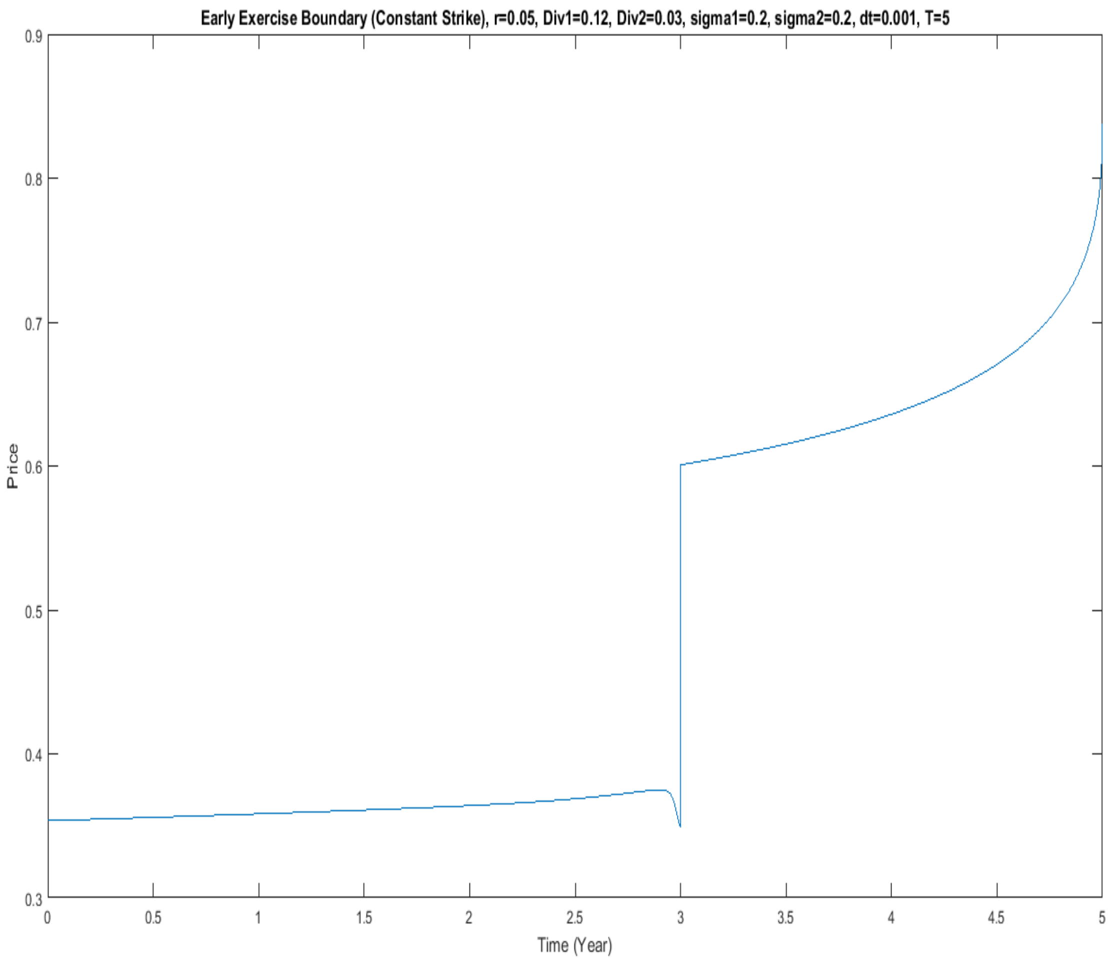

Figure 12: Early exercise boundary for constant strike with regime switching. ($K = 0.8376,\ r = 0.05,\ \delta_1 = 0.12,\ \delta_2 = 0.03,\ \sigma = 0.2,\ dt = 0.001,\ T = 5,\ T_s = 3$)

As a further check of our numerical method, we also compare the numerical and theoretical values of the asymptotic behavior ($t \to -\infty$) of the early exercise boundary. The theoretical values are obtained as follows from the stationary version of the Black-Scholes PDE 3.4.1:

$$S^2 P_{SS} + (k-l) S P_S - kP = 0, \quad where \;\; k = \frac{2r}{\sigma^2}, \;\; l = \frac{2\delta}{\sigma^2} \tag{3.4.16}$$

Suppose $P(S) = aS^b$, we have $P_S(S) = abS^{b-1}$ and $P_{SS}(S) = ab(b-1)S^{b-2}$, and substitution into the asymptotic Equation 3.4.16 gives $b(b-1)+(k-l)b-k = 0$ or $b^2+(k-l-1)b-k = 0$ In order to satisfy the asymptotic condition $P(S) \to 0$ as $S \to +\infty$, we have:

$$b = \frac{-(k-l-1) - \sqrt{(k-l-1)^2 + 4k}}{2} \tag{3.4.17}$$

Denote the asymptotic value as $t \to -\infty$ by S^*. Since $P(S^*) = K - S^*$ and $P_S(S^*) = -1$, we have: $a(S^*)^b = K - S^*$ and $ab(S^*)^{b-1} = -1$. Thus $(1 - \frac{1}{b})S^* = K$, so $S^* = \frac{K}{1-\frac{1}{b}}$.

Case 1: $r = 0.05,\ \delta = 0.12,\ \sigma = 0.2,\ T = 5,\ K = 0.8376,\ b = -0.5$. Thus $S^* = 0.279214857$. See Figure 13 for comparison with value obtained numerically using Equation 3.4.13, integrating back 100 years.

Case 2: $r = 0.05,\ \delta = 0.12,\ \sigma = 0.04577,\ T = 5,\ K = 0.8376,\ b = -0.69660102$. Thus $S^* = 0.34392533$. See Figure 14 for comparison with value obtained numerically, integrating back 100 years.

This, along with the earlier match with existing work [5], suggests that our numerical scheme for solving these generalized Carr integral equations is accurate and robust over long times. Notice that with smaller volatility (case 2) the numerical scheme captures the expected loss of convexity of the early exercise boundary near expiry as discussed in [5].

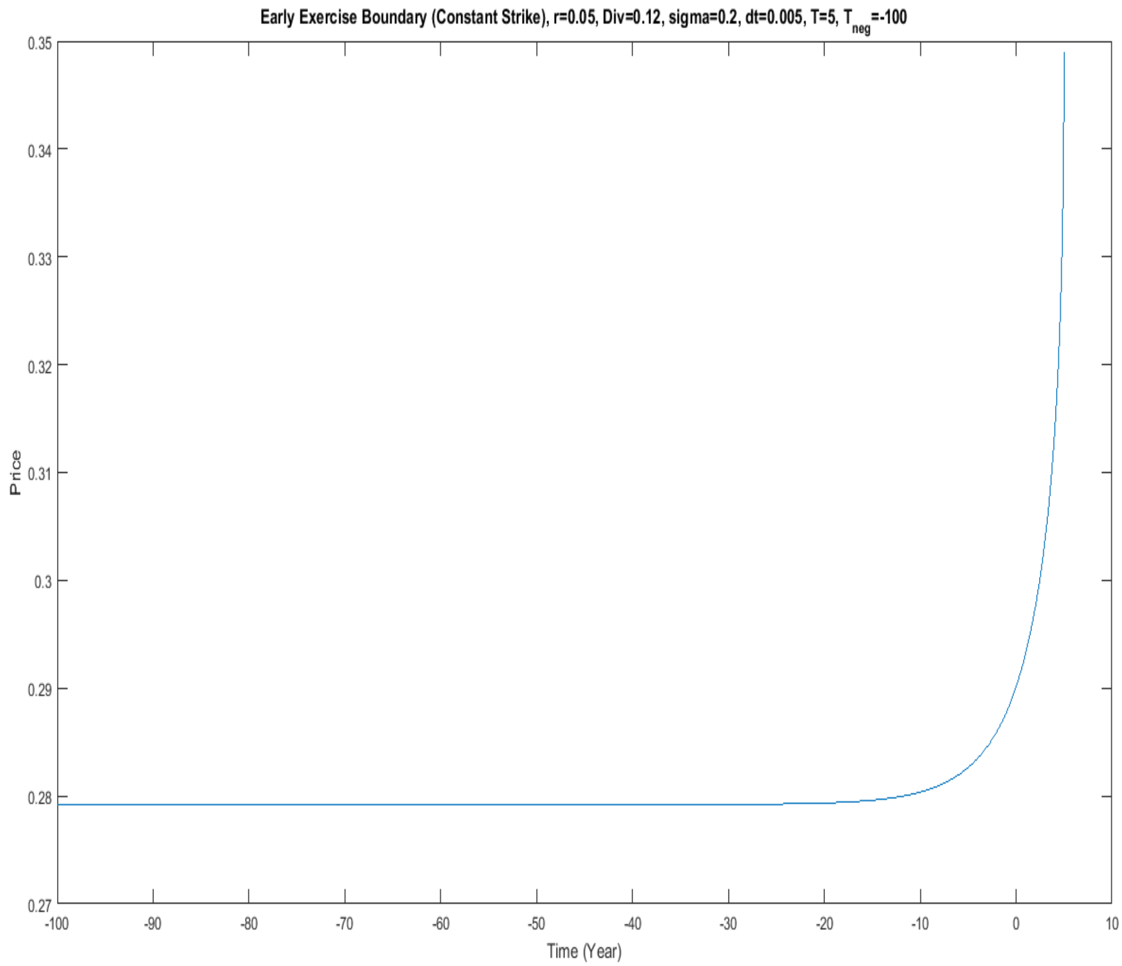

Figure 13: Early exercise boundary for constant strike. ($K = 0.8376,\ r = 0.05,\ \delta = 0.12,\ \sigma = 0.2,\ dt = 0.005,\ T = 5$)

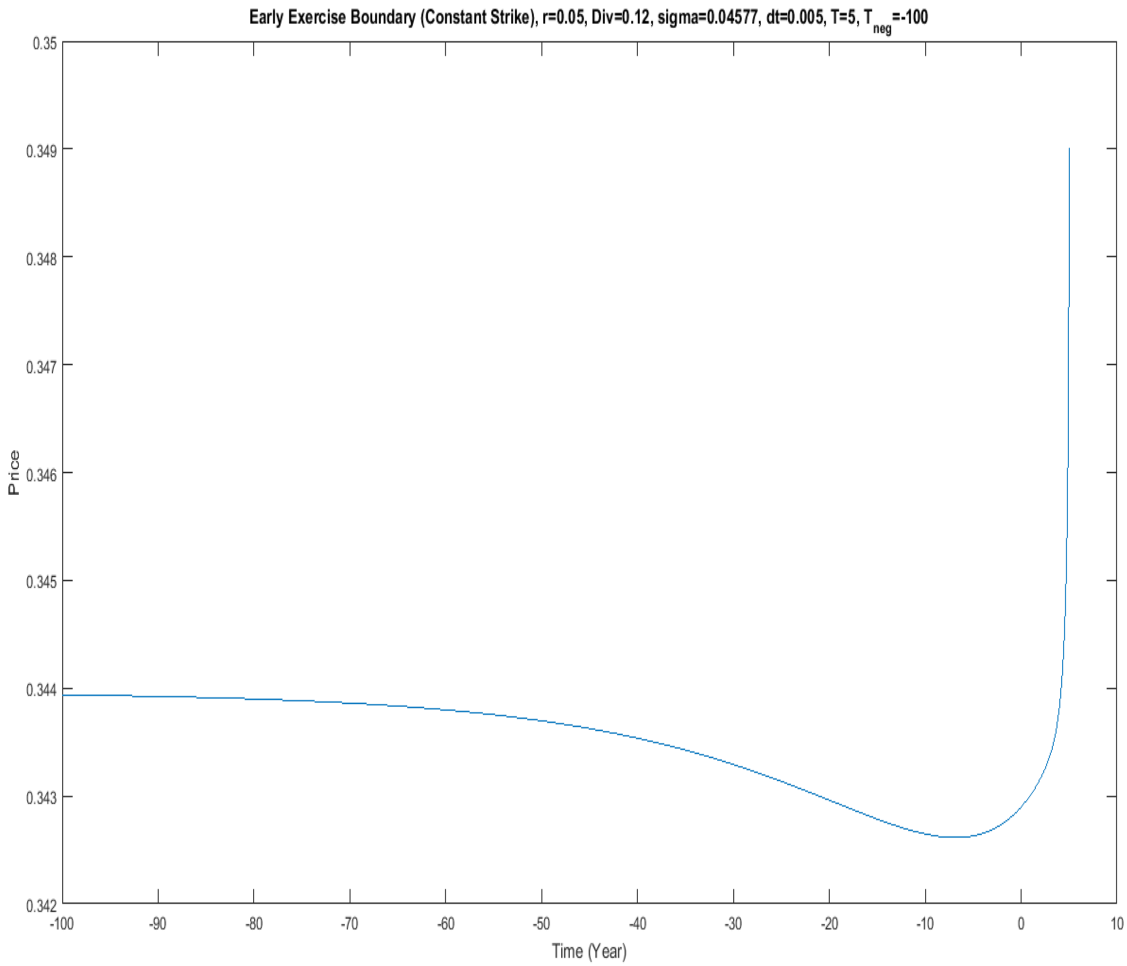

Figure 14: Early exercise boundary for constant strike. $(K = 0.8376,\ r = 0.05,\ \delta = 0.12,\ \sigma = 0.04577,\ dt = 0.005,\ T = 5)$

With this confidence in the numerical scheme, we now apply it to the case where the strike of the option $M(t)$ is a function of t. For the full problem (regime switching, non-constant strike case), we set the parameters as follows: $S(0) = 1$; $T^* = 30$ (Mortgage term, in years); $T = 5$ (Put option maturity, in years); Strike $M(t)$; $T_s = 3$ (Regime switching time, in years); $r = 0.05$; $\delta_1 = 0.12$; $\delta_2 = 0.03$; $\sigma_1 = 0.045768$; $\sigma_2 = 0.031712$; $c = 0.06$; Mortgage To Price Ratio $= 0.9$; $dt = \frac{1}{1000}$ (in year). See Figure 15 for reference. Note that the starting point of the boundary at T_s is $min\left(\frac{r}{\delta_1}M(T_s), B(T_s+)\right)$.

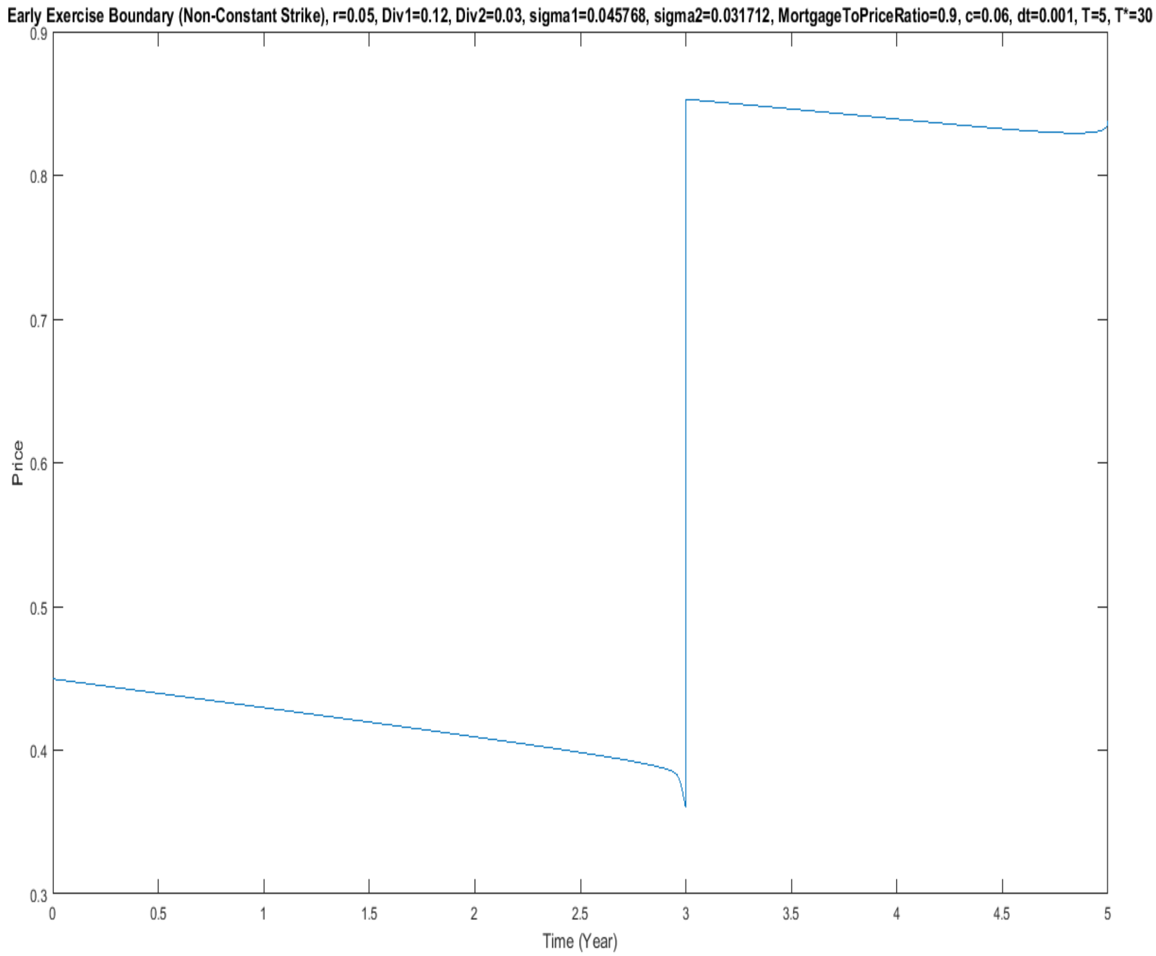

Figure 15: Early exercise boundary for non-constant strike with regime switching. (Strike= $M(t)$, $r = 0.05$, $\delta_1 = 0.12$, $\delta_2 = 0.03$, $\sigma_1 = 0.045768$, $\sigma_2 = 0.031712$, Mortgage To Price Ratio $= 0.9$, $c = 0.06$, $dt = 0.001$, $T = 5$, $T_s = 3$, $T^* = 30$)

We have now completed the first step - we have obtained the early exercise boundary for this American put option with the house as the underlying asset and the remaining value of the mortgage as the strike. Using this early exercise boundary, $B(t)$ in Equation 3.4.6, we now calculate the value of the American put option. In the next section, we shall calculate the probability of exercising this option (the default probability while holding this put), in both the non-regime switching (falling house prices) and the regime switching (falling followed by rising house prices) case.

We first list the parameters that are common to each of the following cases: $M(t)$ as the strike, $S(0) = 1$, Mortgage To Price Ratio = 0.9, $\delta_1 = 0.12$, $\delta_2 = 0.03$, $\sigma_1 = 0.045767598$, $\sigma_2 = 0.031712223$, $\Delta t = 0.005$.

Case 1: $T_s = 3$, $T = 5$.
We find that the price of the American put option
$PutAm(S(0) = 1, t = 0) = 0.024440186223318$. Notice that in the calculation, since we normalized the house price to 1, the option price is approximately 2.44% of the house price. For example, for a \$300,000 house, the option price is \$7,332, which is reasonable for 5-year default protection during a housing slump.

Case 2 (Earlier regime switch from falling to rising house prices): $T_s = 2$, $T = 5$.
We find that the price of the American put option
$PutAm(S(0) = 1, t = 0) = 0.006420005766899$. The price is significantly less since there is (we shall find) reduced chance of foreclosing.

Case 3 (Later regime switch from falling to rising house prices): $T_s = 4$, $T = 5$.
We find that the price of the American put option
$PutAm(S(0) = 1, t = 0) = 0.059429463973766$.

Case 4 (Longer maturity for the American put option): $T_s = 3$, $T = 10$.
We find that the price of the American put option

$PutAm(S(0) = 1, t = 0) = 0.002294106101702$. The price is significantly less than in Case 1 since there is reduced chance of foreclosing in the later years when the house price is rising.

3.4.3 Survival Probability When Holding The Put

As we mentioned earlier, when holding the put the foreclosure does not occur at the first time the house price $S(t)$ drops below the remaining value of the mortgage $M(t)$ as in Section 3.3. Instead, it is optimal to exercise the put option (foreclose) the first time the house price $S(t)$ drops below the early exercise boundary $B(t)$.

We now proceed to derive analytical approximations for the probability of this new type of default. Let $\tau = \inf\{0 \leq t \leq T; S(t) \leq B(t)\}$ and the default probability be $P(\tau < t)$. As in Section 3.3.2, this calculation will be based on a first crossing problem for a Brownian motion which requires linear boundaries for closed form solutions (see Section 3.3.2). Clearly from Figure 15, the early exercise boundary in the full problem (regime switching and with strike $M(t)$) can be fitted quite accurately with two linear curves. Recall that $S(t)$ is a Geometric Brownian Motion, which in phase 1, $0 \leq t \leq T_s$, is give by $S(t) = S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t+\sigma_1 W(t)}$. Suppose that $B(t) \simeq d_1 + m_1 t$ in this interval $[0, T_s]$. We would like to derive the distribution of the default time

$$\tau = \inf\{t > 0, S(t) \leq d_1 + m_1 t\} \tag{3.4.18}$$

The inequality in Equation 3.4.18 is equivalent to

$$W(t) \leq \frac{1}{\sigma_1}\left\{\ln\left(\frac{d_1}{S(0)}\right) - (r - \delta_1 - \frac{\sigma_1^2}{2})t + \ln\left(1 + \frac{m_1}{d_1}t\right)\right\} \tag{3.4.19}$$

As mentioned earlier, closed form solutions are only available for the first crossing of Brownian motions for linear boundaries, so we must make a further approximation of the last term by $\ln\left(1 + \frac{m_1}{d_1}t\right) \approx \frac{m_1}{d_1}t$. Thus we have that $S(t) \leq d_1 + m_1 t$ is approximately equivalent to

$$W(t) \leq \frac{1}{\sigma_1}\left\{\ln\left(\frac{d_1}{S(0)}\right) - \left(r - \delta_1 - \frac{\sigma_1^2}{2} - \frac{m_1}{d_1}\right)t\right\} \tag{3.4.20}$$

Therefore, we may, approximate the survival probability by (see Section 3.3.2)

$$P(\tau > t) = N\left(-\frac{\ln\left(\frac{d_1}{S(0)}\right)}{\sigma_1\sqrt{t}} + \frac{r - \delta_1 - \frac{\sigma_1^2}{2} - \frac{m_1}{d_1}}{\sigma_1}\sqrt{t}\right) - e^{\frac{2(\ln(\frac{d_1}{S(0)}))(r-\delta_1-\frac{\sigma_1^2}{2}-\frac{m_1}{d_1})}{\sigma_1^2}} N\left(\frac{\ln\left(\frac{d_1}{S(0)}\right)}{\sigma_1\sqrt{t}} + \frac{r - \delta_1 - \frac{\sigma_1^2}{2} - \frac{m_1}{d_1}}{\sigma_1}\sqrt{t}\right) \tag{3.4.21}$$

In phase 2, where $T_s \leq t \leq T$, we have, following the same procedure as in Equation 3.3.23 that, with $S(T_s) = y$, the inequality

$$S(t) = ye^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \leq d_2 + m_2 t \tag{3.4.22}$$

is approximately equivalent to

$$W(t - T_s) \leq \frac{1}{\sigma_2}\left\{\ln\left(\frac{d_2}{y}\right) + \frac{m_2}{d_2}T_s - (r - \delta_2 - \frac{\sigma_2^2}{2} - \frac{m_2}{d_2})(t - T_s)\right\} \tag{3.4.23}$$

Using the probability density function for Geometric Brownian Motion [20]:

$$p(S(0), 0; y, t) = \frac{1}{\sigma y\sqrt{2\pi t}} e^{-\frac{\left(\ln\frac{y}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t\right)^2}{2\sigma^2 t}} \tag{3.4.24}$$

we have for $T_s \leq t \leq T$

$$\begin{aligned} &P(\tau > t) \\ &= P(\tau > T_s)\left(\int_{B(T_s-)}^{+\infty} f(y)dy\right)^{-1} \\ &\quad \int_{B(T_s+)}^{+\infty}\left\{N\left(-\frac{b_2}{\sigma_2\sqrt{t-T_s}} + \frac{\mu_2}{\sigma_2}\sqrt{t-T_s}\right) - e^{\frac{2b_2\mu_2}{\sigma_2^2}} N\left(\frac{b_2}{\sigma_2\sqrt{t-T_s}} + \frac{\mu_2}{\sigma_2}\sqrt{t-T_s}\right)\right\} f(y)dy \end{aligned} \tag{3.4.25}$$

where $f(y) = \frac{1}{\sigma_1 y\sqrt{2\pi T_s}} e^{-\frac{\left(\ln\frac{y}{S(0)} - (r-\delta_1-\frac{\sigma_1^2}{2})T_s\right)^2}{2\sigma_1^2 T_s}}$, $b_2 = \ln\left(\frac{d_2}{y}\right) + \frac{m_2}{d_2}T_s$, $\mu_2 = r - \delta_2 - \frac{\sigma_2^2}{2} - \frac{m_2}{d_2}$, and the first term, $P(\tau > T_s)$ is obtained from Equation 3.4.21.

3.4.4 Numerical Results

We now summarize the survival probabilities for various regime switching times. We first list the parameters that are common to each of the following cases:
$M(t)$ as the strike, $S(0) = 1$, Mortgage To Price Ratio $= 0.9$, $\delta_1 = 0.12$, $\delta_2 = 0.03$, $\sigma_1 = 0.045767598$, $\sigma_2 = 0.031712223$, $\Delta t = 0.005$. These are precisely the same as the values chosen to calculate the price of the put in the previous section.

Case 1: $T_s = 3$, $T = 5$
The fitted two linear lines of the boundaries are $y = d_1 + m_1 t = 0.453267902647709 - 0.0196794419974254t$ in phase 1 and $y = d_2 + m_2 t = 0.894155483290446 - 0.0137327534483013t$ in phase 2. The survival probability at time $t = T_s = 3$ is 0.243142604913139 and at $t = T = 5$ is 0.195277070213198. We shall also graph the survival probability for the entire region $0 \leq t \leq T = 5$ in Figure 16 to come in order to compare it with the graph (Figure 9) for the standard model of foreclosure.

Case 2 (Earlier regime switch from falling to rising house prices): $T_s = 2$, $T = 5$.
The fitted two linear lines of the boundaries are $y = d_1 + m_1 t = 0.457171309435733 - 0.0255660274385337t$ in phase 1 and $y = d_2 + m_2 t = 0.893503685576692 - 0.013565982992112t$ in phase 2. The survival probability at time $t = T_s = 2$ is 0.505398931247803 and at $t = T = 5$ is 0.420855783775341.

Case 3 (Later regime switch from falling to rising house prices): $T_s = 4$, $T = 5$
The fitted two linear lines of the boundaries are $y = d_1 + m_1 t = 0.433845003862739 - 0.0130442885119474t$ in phase 1 and $y = d_2 + m_2 t = 0.890854870903941 - 0.0129775183124687t$ in phase 2. The survival probability at time $t = T_s = 4$ is 0.114005244670559 and at $t = T = 5$ is 0.090886183515652.

Case 4 (Longer maturity for the American put option): $T_s = 3$, $T = 10$
The fitted two linear lines of the boundaries are $y = d_1 + m_1 t = 0.545736696123956 -$

$0.0384185962334067t$ in phase 1 and $y = d_2 + m_2 t = 0.909211930566617 - 0.0161231601401344t$ in phase 2. The survival probability at time $t = T = 10$ is 0.165133842276781.

From the numerical results here and in Section 3.4.2, as the regime switching time increases and the other parameters remain the same, the American put option price increases, and the survival probability decreases. Intuitively, as the regime switching time gets longer, the probability of the house price falling below the mortgage balance increases. As a result, the put option is more valuable to the borrower and hence the price of the put increases to compensate the lender for bearing the additional house price risk. The second result evident in the numerical results is that as the American put option maturity increases and other parameters keeping the same, the option price decreases, and the survival probability decreases. In general we see that this optionality from holding the put is a reasonably priced method to avoid adverse credit impacts from foreclosure.

The most significant result arises from comparing the survival probabilities for the simple foreclosure model (Figure 9) with that when holding the put (Figure 16, to come). For this calculation, using Equation 3.4.25, we choose the same parameters as for Figure 9: $T_s = 3$, $T = 5$, $M(t)$ as the strike, $S(0) = 1$, Mortgage To Price Ratio = 0.9, $\delta_1 = 0.12$, $\delta_2 = 0.03$, $\sigma_1 = 0.045767598$, $\sigma_2 = 0.031712223$, $\Delta t = 0.005$.

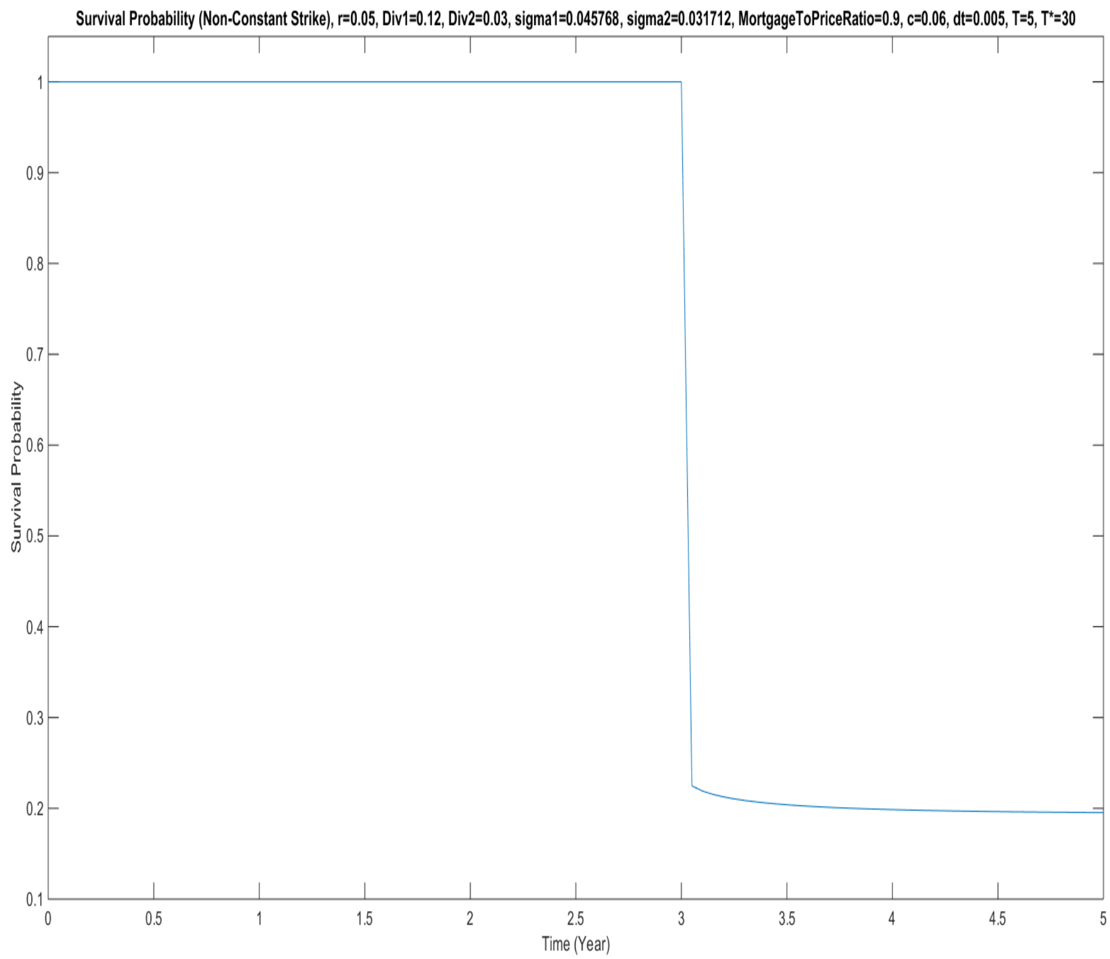

Figure 16: Survival probability for non-constant strike with regime switching. (Strike = $M(t)$, $r = 0.05,\ \delta_1 = 0.12,\ \delta_2 = 0.03,\ \sigma_1 = 0.045768,\ \sigma_2 = 0.031712$, Mortgage To Price Ratio $= 0.9,\ c = 0.06,\ dt = 0.005,\ T = 5,\ T_s = 3,\ T^* = 30$)

Figure 16 shows that if the homeowner follows the optimal strategy for exercising the put, there is essentially no chance of foreclosing as compared to 90% in the standard model summarized in Figure 9. When house prices begin to rise, the homeowner's perception changes to a longer horizon than the maturity of the put ($T = 5$). (S)he knows that eventually the house price will rise above $M(t)$ which is falling to zero at the mortgage maturity, $T^* = 30$.

In summary, we have shown that the put option provides a reasonably priced strategy for avoiding adverse credit impacts to the homeowner should (s)he decide to foreclose during periods of falling house prices. Indeed, comparing Figure 9 and Figure 16, (s)he would almost certainly not foreclose during the period of rapidly falling prices ($0 \leq t \leq T_s = 3$) while holding the put as opposed to an almost certain foreclosure without the put. This reduction in the frequency of foreclosures provides advantages for the mortgage company as well. They would not be forced to hold a non-performing loan or to acquire a devalued house.

3.5 ADJUSTABLE BALANCE MORTGAGE

3.5.1 Theoretical Derivation

At origination, the Adjustable Balance Mortgage (ABM) is like a Fixed-Rate Mortgage (FRM) in that it has a fixed contract rate, has a maturity term and is fully amortizing. If the house value is lower than the originally scheduled balance, the loan balance is set equal to the house value, and the monthly payment is recalculated based on this new value. If the house retains its initial value or increases in value, then the loan balance and payments remain unchanged just as in a standard FRM. Ambrose and Buttimer [2] consider the ABM in discrete time. Here we discuss the ABM in the continuous time case and compare it as an alternative to the American put option discussed in the previous section.

We begin by summarizing the formulas for the standard FRM (see Equation 3.3.8):

$$M(t) = \frac{m}{c}\left(1 - e^{-c(T^*-t)}\right) = \frac{M(0)}{1 - e^{-cT^*}}\left(1 - e^{-c(T^*-t)}\right) \tag{3.5.1}$$

Recall that for an ABM, if the house price is above the remaining value of the mortgage, the mortgage holder pays the mortgage based on the remaining value of the mortgage, as in the FRM case. However, if the house price falls below the remaining value of the mortgage, the mortgage holder's payment is based on the house price, so the payment is lower than the FRM case. Thus for an ABM, Equation 3.5.1 is replaced by

$$\widetilde{M}(t) = \begin{cases} M(t) = \frac{m}{c}\left(1 - e^{-c(T^*-t)}\right) = \frac{M(0)}{1-e^{-cT^*}}\left(1 - e^{-c(T^*-t)}\right) & if\ S(t) \geq M(t) \\ S(t) \quad if\ S(t) < M(t) \end{cases} \tag{3.5.2}$$

with the adjusted mortgage payment being

$$\widetilde{m}(t) = \begin{cases} m & if\ S(t) \geq M(t) \\ a_t & if\ S(t) < M(t) \end{cases} \tag{3.5.3}$$

where a_t is the annualized mortgage payment calculated on the house price at time t (i.e. $a_t dt$ is paid in $[t, t+dt]$). Note that a_t is a constant during $[t, t+dt]$ but it keeps changing as $S(t)$ changes. Comparing with Equation 3.5.1, we have:

$$S(t) = \frac{a_t}{c}\left(1 - e^{-c(T^*-t)}\right) \tag{3.5.4}$$

So

$$a_t = \frac{S(t)c}{1 - e^{-c(T^*-t)}} \tag{3.5.5}$$

In order to compare the upfront values of the FRM and ABM, we calculate the discounted cash flows between 0 to T, discounting at the rate c, of a T^* (*say* 30) year. For the FRM this (deterministic) value is

$$\int_0^T e^{-ct} m dt = m \int_0^T e^{-ct} dt = \frac{m}{c}\left(1 - e^{-cT}\right) = \frac{M(0)}{1 - e^{-cT^*}}\left(1 - e^{-cT}\right) \tag{3.5.6}$$

The corresponding calculation for the ABM is

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}\left[\int_0^T e^{-ct} d\widetilde{m}(t)\right] &= \mathbb{E}_{\mathbb{P}}\left[\int_0^T e^{-ct}\left\{m\chi_{[S(t)\geq M(t)]} + a_t\chi_{[S(t)<M(t)]}\right\} dt\right] \\ &= m\int_0^T e^{-ct}\mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)\geq M(t)]}\right] dt + \int_0^T e^{-ct}\mathbb{E}_{\mathbb{P}}\left[a_t\chi_{[S(t)<M(t)]}\right] dt \\ &= m\int_0^T e^{-ct} P\left(S(t) \geq M(t)\right) dt + \int_0^T \frac{ce^{-ct}}{1 - e^{-c(T^*-t)}}\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] dt \end{aligned} \tag{3.5.7}$$

The difference between Equation 3.5.6 and Equation 3.5.7 is the price that the mortgage company would charge for this optionality.

i) Non-Regime Switching Case:

We first do the calculations in Equation 3.5.7 in the simple case with no regime switching. As in Equation 3.3.14, the inequality

$S(t) = S(0)e^{(r-\delta-\frac{\sigma^2}{2})t+\sigma W(t)} \geq M(t) = \frac{M(0)}{1-e^{-cT^*}}\left(1-e^{-c(T^*-t)}\right)$ is equivalent to

$$W(t) \geq \frac{1}{\sigma}\left\{\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta-\frac{\sigma^2}{2}\right)t\right\} \tag{3.5.8}$$

Since $W(t) \sim N(0,t)$, we have:

$$\begin{aligned}
&P(S(t) \geq M(t)) \\
&= P\left(W(t) \geq \frac{1}{\sigma}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta-\frac{\sigma^2}{2}\right)t\right]\right) \\
&= 1 - P\left(W(t) \leq \frac{1}{\sigma}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta-\frac{\sigma^2}{2}\right)t\right]\right) \\
&= 1 - \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\frac{1}{\sigma}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)}+\ln\left(1-e^{-c(T^*-t)}\right)-\left(r-\delta-\frac{\sigma^2}{2}\right)t\right]} e^{-\frac{x^2}{2t}}dx
\end{aligned} \tag{3.5.9}$$

Let $z = \frac{x}{\sqrt{t}}$, then $dx = \sqrt{t}dz$, and Equation 3.5.9 becomes

$$\begin{aligned}
&P(S(t) \geq M(t)) \\
&= 1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{1}{\sigma\sqrt{t}}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)}+\ln\left(1-e^{-c(T^*-t)}\right)-\left(r-\delta-\frac{\sigma^2}{2}\right)t\right]} e^{-\frac{z^2}{2}}dz \\
&= 1 - N\left(\frac{1}{\sigma\sqrt{t}}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta-\frac{\sigma^2}{2}\right)t\right]\right)
\end{aligned} \tag{3.5.10}$$

Next we calculate $\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right]$ using the known transition probability density function for a Geometric Brownian Motion [20].

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] &= \int_0^{+\infty} y\chi_{[y<M(t)]}\frac{1}{\sigma y\sqrt{2\pi t}}e^{-\frac{\left(\ln\frac{y}{S(0)}-(r-\delta-\frac{\sigma^2}{2})t\right)^2}{2\sigma^2 t}}dy \\
&= \frac{1}{\sigma\sqrt{2\pi t}}\int_0^{M(t)} e^{-\frac{\left(\ln\frac{y}{S(0)}-(r-\delta-\frac{\sigma^2}{2})t\right)^2}{2\sigma^2 t}}dy
\end{aligned} \tag{3.5.11}$$

Let $z = \frac{\ln \frac{y}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$, then $dz = \frac{1}{\sigma\sqrt{t}}\frac{1}{y}dy$, and with $y = S(0)e^{\sigma\sqrt{t}z + (r-\delta-\frac{\sigma^2}{2})t}$, we have:

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{\ln\frac{M(t)}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}} ye^{-\frac{z^2}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{(r-\delta-\frac{\sigma^2}{2})t}\int_{-\infty}^{\frac{\ln\frac{M(t)}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}} e^{\sigma\sqrt{t}z-\frac{z^2}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{(r-\delta-\frac{\sigma^2}{2})t}\int_{-\infty}^{\frac{\ln\frac{M(t)}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}} e^{-\frac{(z-\sigma\sqrt{t})^2}{2}}e^{\frac{\sigma^2 t}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{(r-\delta)t}\int_{-\infty}^{\frac{\ln\frac{M(t)}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}} e^{-\frac{(z-\sigma\sqrt{t})^2}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{(r-\delta)t}\int_{-\infty}^{\frac{\ln\frac{M(t)}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}-\sigma\sqrt{t}} e^{-\frac{q^2}{2}}dq \\
&= S(0)e^{(r-\delta)t}N\left(\frac{\ln\frac{M(t)}{S(0)} - (r-\delta-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}} - \sigma\sqrt{t}\right) \\
&= S(0)e^{(r-\delta)t}N\left(\frac{\ln\frac{M(t)}{S(0)} - (r-\delta+\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right)
\end{aligned} \tag{3.5.12}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[\int_0^T e^{-ct}d\widetilde{m}(t)\right] \\
&= m\int_0^T e^{-ct}P\left(S(t)\geq M(t)\right)dt + \int_0^T \frac{ce^{-ct}}{1-e^{-c(T^*-t)}}\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right]dt \\
&= m\int_0^T e^{-ct}\left\{1 - N\left(\frac{1}{\sigma\sqrt{t}}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta-\frac{\sigma^2}{2}\right)t\right]\right)\right\} \\
&\qquad dt \\
&\quad + cS(0)\int_0^T \frac{e^{(r-\delta-c)t}}{1-e^{-c(T^*-t)}}N\left(\frac{\ln\frac{M(t)}{S(0)} - (r-\delta+\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right)dt
\end{aligned} \tag{3.5.13}$$

ii) Regime Switching Case:

In this case, the expression corresponding to Equation 3.5.7 is given by

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[\int_0^T e^{-ct} d\widetilde{m}(t)\right] = \mathbb{E}_{\mathbb{P}}\left[\int_0^T e^{-ct}\left\{m\chi_{[S(t)\geq M(t)]} + a_t\chi_{[S(t)<M(t)]}\right\} dt\right] \\
&= \int_0^{T_s} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)\geq M(t)]}\right] mdt + \int_{T_s}^{T} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)\geq M(t)]}\right] mdt \\
&\quad + \int_0^{T_s} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[a_t\chi_{[S(t)<M(t)]}\right] dt + \int_{T_s}^{T} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[a_t\chi_{[S(t)<M(t)]}\right] dt \\
&= m\int_0^{T_s} e^{-ct}P\left(S(t)\geq M(t)\right) dt + m\int_{T_s}^{T} e^{-ct}P\left(S(t)\geq M(t)\right) dt \\
&\quad + \int_0^{T_s} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[a_t\chi_{[S(t)<M(t)]}\right] dt + \int_{T_s}^{T} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[a_t\chi_{[S(t)<M(t)]}\right] dt \\
&= part1 + part2 + part3 + part4
\end{aligned} \tag{3.5.14}$$

We now calculate each of the four parts separately.

For $part1 = m\int_0^{T_s} e^{-ct}P\left(S(t)\geq M(t)\right) dt$, $P(S(t)\geq M(t))$ is simply a repeat of Equation 3.5.10 with $\delta = \delta_1$ and $\sigma = \sigma_1$.

$$\begin{aligned}
&P(S(t)\geq M(t)) \\
&= 1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{1}{\sigma_1\sqrt{t}}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)t\right]} e^{-\frac{z^2}{2}} dz \\
&= 1 - N\left(\frac{1}{\sigma_1\sqrt{t}}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)t\right]\right)
\end{aligned} \tag{3.5.15}$$

Therefore,

$$\begin{aligned}
&part1 = m\int_0^{T_s} e^{-ct}P\left(S(t)\geq M(t)\right) dt \\
&= m\int_0^{T_s} e^{-ct} \\
&\qquad \left\{1 - N\left(\frac{1}{\sigma_1\sqrt{t}}\left[\ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)t\right]\right)\right\} dt
\end{aligned} \tag{3.5.16}$$

For $part2 = m\int_{T_s}^{T} e^{-ct}P\left(S(t) \geq M(t)\right)dt$, we begin by calculating $P\left(S(t) \geq M(t)\right)$ when $T_s \leq t \leq T$.

$$\begin{aligned} S(t) &= S(T_s)e^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \\ &= S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+\sigma_1 W(T_s)}e^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \\ &= S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+\sigma_1 W(T_s)+(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \end{aligned} \tag{3.5.17}$$

So $S(t) \geq M(t)$ is equivalent to

$$S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+\sigma_1 W(T_s)+(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \geq \frac{M(0)}{1-e^{-cT^*}}\left(1-e^{-c(T^*-t)}\right) \tag{3.5.18}$$

which is equivalent to

$$\begin{aligned} \sigma_1 W(T_s)+\sigma_2 W(t-T_s) \geq \ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s \\ - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s) \end{aligned} \tag{3.5.19}$$

For Brownian motion, $W(T_s)$ and $W(t-T_s)$ are independent, since the time intervals $[0, T_s]$ and $[T_s, t]$ are non-overlapping for $T_s \leq t \leq T$. We also know that $W(T_s) \sim N(0, T_s)$ and $W(t-T_s) \sim N(0, t-T_s)$, so $\sigma_1 W(T_s) + \sigma_2 W(t-T_s) \sim N(0, \sigma_1^2 T_s + \sigma_2^2(t-T_s))$. Thus

$$\begin{aligned} &P(S(t) \geq M(t)) \\ &= P(\sigma_1 W(T_s)+\sigma_2 W(t-T_s) \geq \ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) \\ &\qquad - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)) \\ &= 1 - P(\sigma_1 W(T_s)+\sigma_2 W(t-T_s) \leq \ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) \\ &\qquad - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)) \\ &= 1 - \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma_1^2 T_s+\sigma_2^2(t-T_s)}}\int_{-\infty}^{UpperLimit_1} e^{-\frac{x^2}{2(\sigma_1^2 T_s+\sigma_2^2(t-T_s))}}dx \end{aligned} \tag{3.5.20}$$

where

$UpperLimit_1 = \ln\frac{M(0)}{S(0)\left(1-e^{-cT^*}\right)} + \ln\left(1-e^{-c(T^*-t)}\right) - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)$

Let $z = \frac{x}{\sqrt{\sigma_1^2 T_s + \sigma_2^2 (t - T_s)}}$, then $dx = \sqrt{\sigma_1^2 T_s + \sigma_2^2 (t - T_s)} dz$. Then

$$P(S(t) \geq M(t)) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{UpperLimit_2} e^{-\frac{z^2}{2}} dz = 1 - N\left(UpperLimit_2\right) \tag{3.5.21}$$

where

$UpperLimit_2 = \frac{1}{\sqrt{\sigma_1^2 T_s + \sigma_2^2 (t - T_s)}} \times$
$\left[\ln \frac{M(0)}{S(0)\left(1 - e^{-cT^*}\right)} + \ln\left(1 - e^{-c(T^* - t)}\right) - \left(r - \delta_1 - \frac{\sigma_1^2}{2}\right) T_s - \left(r - \delta_2 - \frac{\sigma_2^2}{2}\right)(t - T_s)\right]$

Therefore,

$$part2 = m \int_{T_s}^{T} e^{-ct} P\left(S(t) \geq M(t)\right) dt = m \int_{T_s}^{T} e^{-ct}\left(1 - N\left(UpperLimit_2\right)\right) dt \tag{3.5.22}$$

For $part3$, $\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right]$ is obtained from Equation 3.5.12 with $\delta = \delta_1$ and $\sigma = \sigma_1$. Therefore,

$$\begin{aligned} part3 &= \int_0^{T_s} e^{-ct} \mathbb{E}_{\mathbb{P}}\left[a_t \chi_{[S(t)<M(t)]}\right] dt = \int_0^{T_s} e^{-ct} \mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)<M(t)]} \frac{cS(t)}{1 - e^{-c(T^* - t)}}\right] dt \\ &= \int_0^{T_s} e^{-ct} \frac{c}{1 - e^{-c(T^* - t)}} \mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] dt \\ &= cS(0) \int_0^{T_s} \frac{e^{(r - \delta_1 - c)t}}{1 - e^{-c(T^* - t)}} N\left(\frac{\ln \frac{M(t)}{S(0)} - (r - \delta_1 + \frac{\sigma_1^2}{2})t}{\sigma_1 \sqrt{t}}\right) dt \end{aligned} \tag{3.5.23}$$

For $part4$, we have:

$$\begin{aligned} part4 &= \int_{T_s}^{T} e^{-ct} \mathbb{E}_{\mathbb{P}}\left[a_t \chi_{[S(t)<M(t)]}\right] dt = \int_{T_s}^{T} e^{-ct} \mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)<M(t)]} \frac{cS(t)}{1 - e^{-c(T^* - t)}}\right] dt \\ &= \int_{T_s}^{T} e^{-ct} \frac{c}{1 - e^{-c(T^* - t)}} \mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] dt \end{aligned} \tag{3.5.24}$$

In order to calculate $\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right]$, we must first calculate the probability density function for $S(t)$ on $[T_s, T]$.

$$P(S(t) \le y) = P(S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+\sigma_1 W(T_s)+(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \le y)$$
$$= P\left(\sigma_1 W(T_s) + \sigma_2 W(t-T_s) \le \ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right) \tag{3.5.25}$$

Since $\sigma_1 W(T_s) + \sigma_2 W(t-T_s) \sim N(0, \sigma_1^2 T_s + \sigma_2^2(t-T_s))$, we have:

$$P(S(t) \le y)$$
$$= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma_1^2 T_s + \sigma_2^2(t-T_s)}}\int_{-\infty}^{\ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)} e^{-\frac{x^2}{2(\sigma_1^2 T_s+\sigma_2^2(t-T_s))}}dx \tag{3.5.26}$$

Let $z = \frac{x}{\sqrt{\sigma_1^2 T_s+\sigma_2^2(t-T_s)}}$, then $dx = \sqrt{\sigma_1^2 T_s + \sigma_2^2(t-T_s)}dz$, so:

$$P(S(t) \le y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{1}{\sqrt{\sigma_1^2 T_s+\sigma_2^2(t-T_s)}}\left[\ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right]} e^{-\frac{z^2}{2}}dz$$
$$= N\left(\frac{1}{\sqrt{\sigma_1^2 T_s + \sigma_2^2(t-T_s)}}\left[\ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right]\right) \tag{3.5.27}$$

Thus the probability density function for $S(t)$ on $[T_s, T]$ is:

$$\frac{d}{dy}P(S(t) \le y)$$
$$= \frac{d}{dy}N\left(\frac{1}{\sqrt{\sigma_1^2 T_s + \sigma_2^2(t-T_s)}}\left[\ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right]\right)$$
$$= \frac{1}{\sqrt{2\pi}}e^{-\frac{\left(\ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right)^2}{2(\sigma_1^2 T_s+\sigma_2^2(t-T_s))}}\frac{1}{\sqrt{\sigma_1^2 T_s + \sigma_2^2(t-T_s)}}\frac{1}{y}$$
$$= \frac{1}{y\sqrt{2\pi}\sqrt{\sigma_1^2 T_s + \sigma_2^2(t-T_s)}}e^{-\frac{\left(\ln\frac{y}{S(0)} - \left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s - \left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right)^2}{2(\sigma_1^2 T_s+\sigma_2^2(t-T_s))}} \tag{3.5.28}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] \\
&= \int_0^{+\infty} y\chi_{[y<M(t)]}\frac{1}{y\sqrt{2\pi}\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}}e^{-\frac{\left(\ln\frac{y}{S(0)}-\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s-\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right)^2}{2(\sigma_1^2T_s+\sigma_2^2(t-T_s))}}dy \\
&= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}}\int_0^{M(t)} e^{-\frac{\left(\ln\frac{y}{S(0)}-\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s-\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)\right)^2}{2(\sigma_1^2T_s+\sigma_2^2(t-T_s))}}dy
\end{aligned} \tag{3.5.29}$$

Let $z = \frac{\ln\frac{y}{S(0)}-\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s-\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)}{\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}}$, then $dz = \frac{1}{\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}}\frac{1}{y}dy$, and with $y = S(0)e^{\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}z+\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s+\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)}$, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{UpperLimit_3} ye^{-\frac{z^2}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s+\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)}\int_{-\infty}^{UpperLimit_3} e^{\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}z-\frac{z^2}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s+\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)+\frac{1}{2}(\sigma_1^2T_s+\sigma_2^2(t-T_s))}\int_{-\infty}^{UpperLimit_3} e^{-\frac{\left(z-\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}\right)^2}{2}}dz \\
&= \frac{S(0)}{\sqrt{2\pi}}e^{\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s+\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)+\frac{1}{2}(\sigma_1^2T_s+\sigma_2^2(t-T_s))}\int_{-\infty}^{UpperLimit_4} e^{-\frac{q^2}{2}}dq \\
&= S(0)e^{\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s+\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)+\frac{1}{2}(\sigma_1^2T_s+\sigma_2^2(t-T_s))}N\left(UpperLimit_4\right)
\end{aligned} \tag{3.5.30}$$

where

$$UpperLimit_3 = \frac{\ln\frac{M(t)}{S(0)}-\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s-\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)}{\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}}$$

$$UpperLimit_4 = \frac{\ln\frac{M(t)}{S(0)}-\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s-\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)}{\sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}} - \sqrt{\sigma_1^2T_s+\sigma_2^2(t-T_s)}$$

Therefore,

$$
\begin{aligned}
&part4 \\
&= \int_{T_s}^{T} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)<M(t)]}da(t)\right] = \int_{T_s}^{T} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)<M(t)]}a_t dt\right] \\
&= \int_{T_s}^{T} e^{-ct}\mathbb{E}_{\mathbb{P}}\left[\chi_{[S(t)<M(t)]}\frac{cS(t)}{1-e^{-c(T^*-t)}}\right]dt = \int_{T_s}^{T} e^{-ct}\frac{c}{1-e^{-c(T^*-t)}}\mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right]dt \\
&= cS(0)\int_{T_s}^{T}\frac{e^{-ct}}{1-e^{-c(T^*-t)}}e^{\left(r-\delta_1-\frac{\sigma_1^2}{2}\right)T_s+\left(r-\delta_2-\frac{\sigma_2^2}{2}\right)(t-T_s)+\frac{1}{2}(\sigma_1^2T_s+\sigma_2^2(t-T_s))}N\left(UpperLimit_4\right)dt
\end{aligned}
\tag{3.5.31}
$$

In the next section we will use these expressions to calculate the upfront additional payment for the ABM over the traditional FRM.

We now calculate the expectation of the ABM modified payment stream a_t.

Since

$$a_t = \frac{S(t)c}{1-e^{-c(T^*-t)}} \tag{3.5.32}$$

then

$$E\left[a_t\right] = \frac{c}{1-e^{-c(T^*-t)}}E\left[S(t)\right] \tag{3.5.33}$$

In the regime switching case when $0 \le t \le T_s$,

$$S(t) = S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t+\sigma_1 W(t)} \tag{3.5.34}$$

So

$$E\left[S(t)\right] = S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t}E\left[e^{\sigma_1 W(t)}\right] \tag{3.5.35}$$

We know that for a random variable $X \sim N(\mu,\sigma^2)$, the moment-generating function $M_X(u) = E\left[e^{uX}\right] = e^{\mu u+\frac{1}{2}\sigma^2u^2}$.

Since $\sigma_1 W(t) \sim N(0,\sigma_1^2 t)$ we have

$$E\left[e^{\sigma_1 W(t)}\right] = E\left[e^{u\sigma_1 W(t)}\right]\Big|_{u=1} = e^{\frac{1}{2}\sigma_1^2tu^2}\Big|_{u=1} = e^{\frac{1}{2}\sigma_1^2t} \tag{3.5.36}$$

Thus

$$E\left[S(t)\right] = S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t}E\left[e^{\sigma_1 W(t)}\right] = S(0)e^{(r-\delta_1-\frac{\sigma_1^2}{2})t}e^{\frac{1}{2}\sigma_1^2t} = S(0)e^{(r-\delta_1)t} \tag{3.5.37}$$

Therefore

$$E[a_t] = \frac{cS(0)}{1 - e^{-c(T^*-t)}} e^{(r-\delta_1)t} \quad when\ 0 \le t \le T_s \tag{3.5.38}$$

When $T_s \le t \le T$,

$$\begin{aligned} S(t) &= S(T_s) e^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \\ &= S(0) e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+\sigma_1 W(T_s)} e^{(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_2 W(t-T_s)} \\ &= S(0) e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\sigma_1 W(T_s)+\sigma_2 W(t-T_s)} \end{aligned} \tag{3.5.39}$$

Thus

$$E[S(t)] = S(0) e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)} E\left[e^{\sigma_1 W(T_s)+\sigma_2 W(t-T_s)}\right] \tag{3.5.40}$$

As discussed earlier, we know that $X := \sigma_1 W(T_s) + \sigma_2 W(t - T_s) \sim N(0, \sigma_1^2 T_s + \sigma_2^2(t - T_s))$ and hence

$$E\left[e^{\sigma_1 W(T_s)+\sigma_2 W(t-T_s)}\right] = E\left[e^{uX}\right]\Big|_{u=1} = e^{\frac{1}{2}(\sigma_1^2 T_s+\sigma_2^2(t-T_s))} \tag{3.5.41}$$

Thus

$$\begin{aligned} E[a_t] &= \frac{c}{1 - e^{-c(T^*-t)}} E[S(t)] \\ &= \frac{cS(0)}{1 - e^{-c(T^*-t)}} e^{(r-\delta_1-\frac{\sigma_1^2}{2})T_s+(r-\delta_2-\frac{\sigma_2^2}{2})(t-T_s)+\frac{1}{2}(\sigma_1^2 T_s+\sigma_2^2(t-T_s))} \quad when\ T_s \le t \le T \end{aligned} \tag{3.5.42}$$

We also note that the probability distribution for the time when the ABM is modified, $P(S(t) \le M(t))$, has already been calculated in Equation 3.5.15 and Equation 3.5.21.

3.5.2 Numerical Results

In order to compare the ABM optionality with the American put option, we list the ABM results here, in both the non-regime switching case and the regime switching cases. In each of the ABM cases, we compare them with the corresponding FRM result. First, we will list the FRM result.

3.5.2.1 FRM

We calculate the discounted value of a T^* year FRM cash flow between 0 to T ($< T^*$) years, discounting at rate c. Using Equation 3.5.6:

$$PV(0)_{FRM} = \int_0^T e^{-ct} m dt = \frac{M(0)}{1 - e^{-cT^*}} \left(1 - e^{-cT}\right) \tag{3.5.43}$$

with the parameters $S(0) = 1$; $T^* = 30$ (Mortgage term, in years); $T = 5$ (Maturity, in years); $r = 0.05$; $c = 0.06$; Mortgage To Price Ratio $= 0.9$, we obtain

$$PV(0)_{FRM} = 0.279457638302964 \tag{3.5.44}$$

3.5.2.2 ABM

i) Non-Regime Switching Case

Using Equation 3.5.13

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}}\left[\int_0^T e^{-ct} d\widetilde{m}(t)\right] \\
&= m \int_0^T e^{-ct} P\left(S(t) \geq M(t)\right) dt + \int_0^T \frac{ce^{-ct}}{1 - e^{-c(T^*-t)}} \mathbb{E}_{\mathbb{P}}\left[S(t)\chi_{[S(t)<M(t)]}\right] dt \\
&= m \int_0^T e^{-ct} \left\{1 - N\left(\frac{1}{\sigma\sqrt{t}}\left[\ln \frac{M(0)}{S(0)\left(1 - e^{-cT^*}\right)} + \ln\left(1 - e^{-c(T^*-t)}\right) - \left(r - \delta - \frac{\sigma^2}{2}\right) t\right]\right)\right\} \\
&\qquad dt \\
&\quad + cS(0) \int_0^T \frac{e^{(r-\delta-c)t}}{1 - e^{-c(T^*-t)}} N\left(\frac{\ln \frac{M(t)}{S(0)} - (r - \delta + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\right) dt
\end{aligned} \tag{3.5.45}$$

With the parameters above augmented with those for the house price in phase 1: $S(0) = 1$; $T^* = 30$ (Mortgage term, in years); $T = 5$ (Maturity, in years); $r = 0.05$; $\delta = 0.12$; $\sigma = 0.045767598$; $c = 0.06$; Mortgage To Price Ratio$= 0.9$, we obtain

$$PV(0)_{ABM} = 0.263975139455276 \tag{3.5.46}$$

Note that the result is less than for the FRM, as expected, since the house price is decreasing ($\mu = r - \delta = -0.07$). A person should pay $0.279457638302964 - 0.263975139455276 = 0.015482498847688$ percent of the house price upfront to hold this option during a housing decline.

In order to visualize the ABM present value with respect to the dividend rate ($r - \delta$ is the rate of decline of house price), we plot Figure 17. As expected, when the house price begins to decline the ABM price begins to differ from the FRM price.

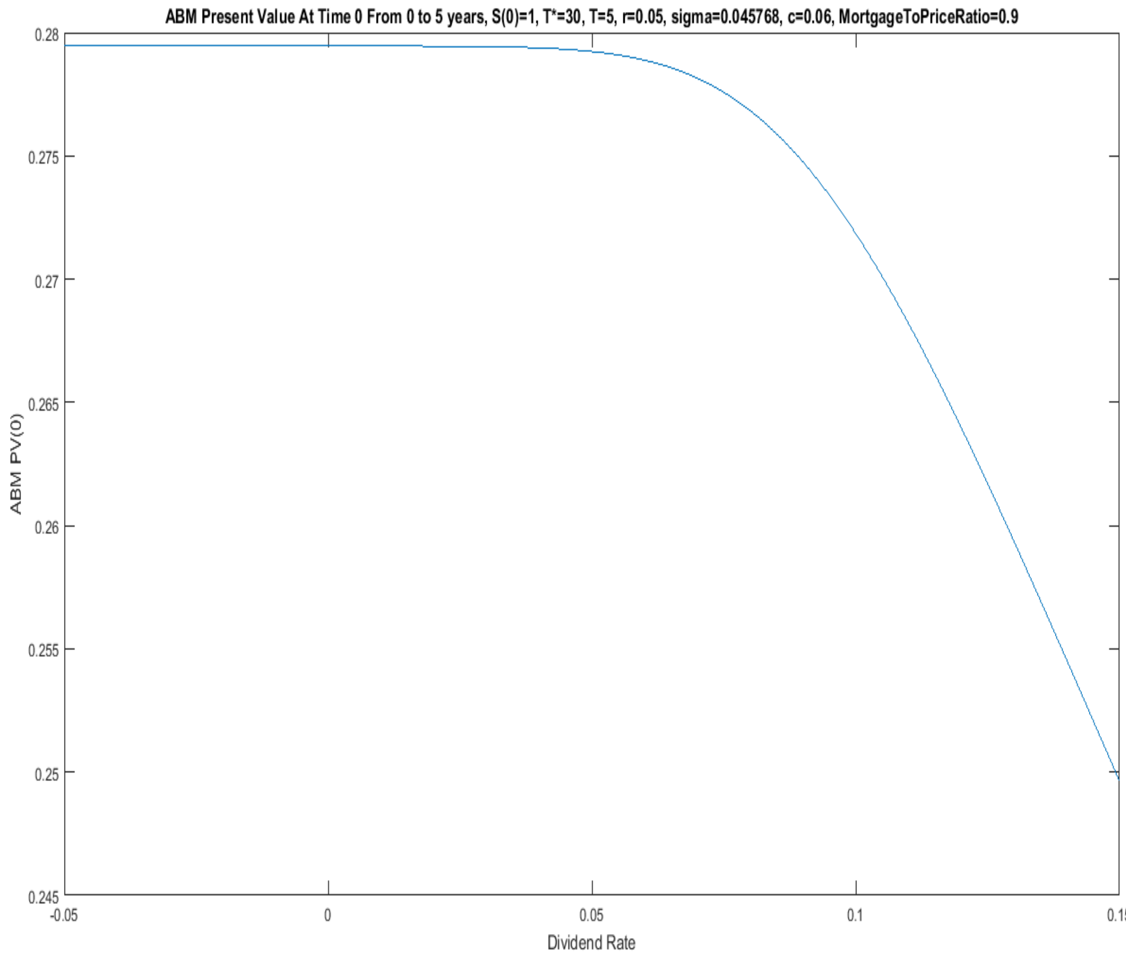

Figure 17: ABM Present Value At Time 0 From 0 to 5 years, Non-RS Case. ($S(0) = 1,\ T^* = 30,\ T = 5,\ r = 0.05,\ \sigma = 0.045768,\ c = 0.06$, Mortgage To Price Ratio $= 0.9$)

ii) Regime Switching Case

In order to compare the ABM present value in the non-regime switching and regime switching cases, we simultaneously plot in Figure 18 the ABM present values for variable δ_1 throughout $0 \leq t \leq T = 5$ (blue curve as in Figure 17) and with variable δ_1 in phase 1, $0 \leq t \leq T_s = 3$ followed by $\delta_2 = 0.03$ for rising prices in phase 2, $T_s \leq t \leq T$ (red curve). As shown in the graph, the regime switching ABM present value is always equal to or higher than that in the non-regime switching case, due to the house price increase in phase 2.

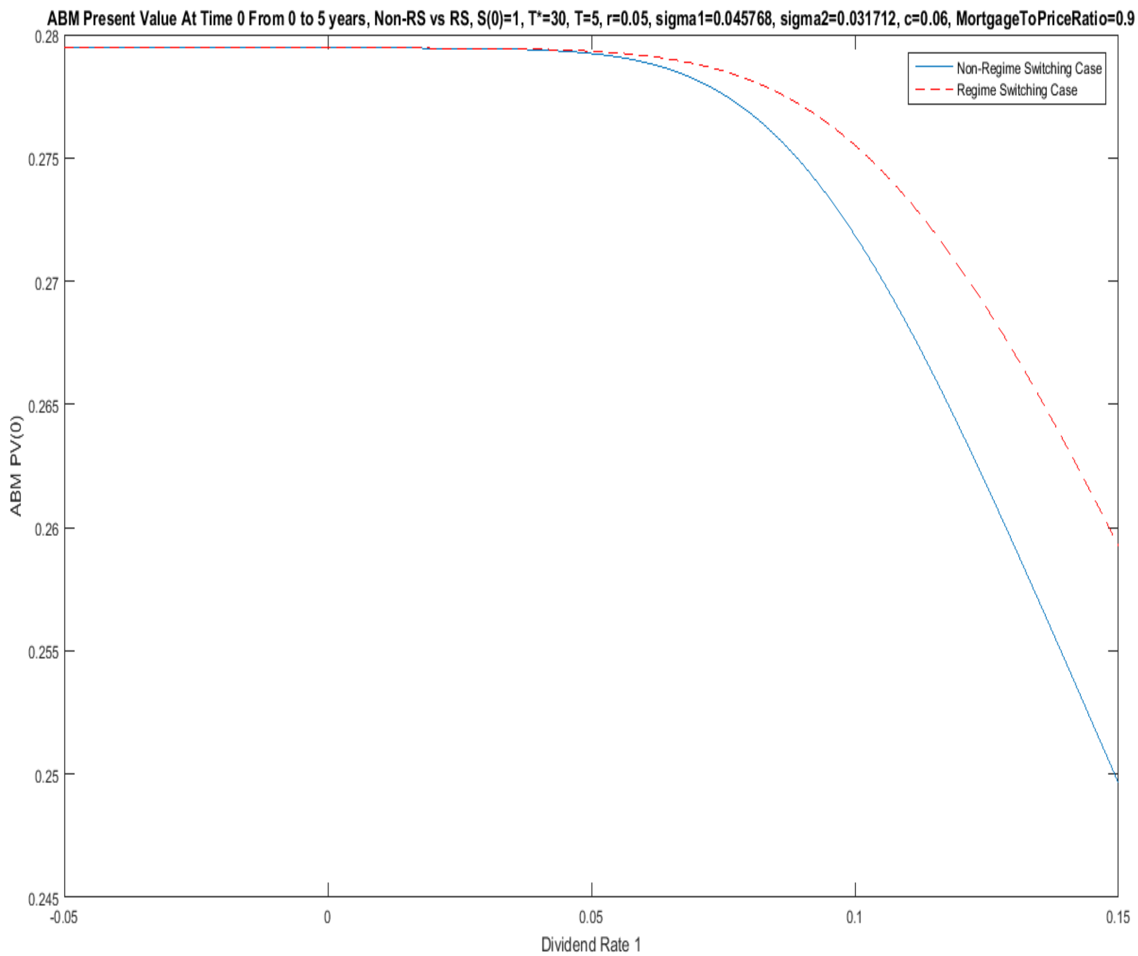

Figure 18: ABM Present Value At Time 0 From 0 to 5 years, Non-RS vs RS. ($S(0) = 1,\ T^* = 30,\ T = 5,\ r = 0.05,\ \sigma_1 = 0.045768,\ \sigma_2 = 0.031712,\ \delta_2 = 0.03,\ c = 0.06$, Mortgage To Price Ratio $= 0.9$)

In order to compare the ABM strategy with the American put option case, we pick a similar set of parameters in four cases as follows.

We list together here the parameters that are common to each of the following cases: $S(0) = 1$; $T^* = 30$ (Mortgage term, in years); $r = 0.05$; $\delta_1 = 0.12$; $\delta_2 = 0.03$; $\sigma_1 = 0.045767598$; $\sigma_2 = 0.031712223$; $c = 0.06$; Mortgage To Price Ratio $= 0.9$;

Case 1: $T_s = 3$, $T = 5$

$$PV(0)_{ABMRS} = 0.270517288796984 \tag{3.5.47}$$

Note that the result is higher than the non-regime switching case since house prices rise in the last 2 years. A person should pay $0.279457638302964 - 0.270517288796984 = 0.00894034950598$ percent of the house price upfront for this option. If the mortgage company received 0.015482498847688 (based on assuming falling house prices for the entire $T = 5$ years) and this regime switching happens, the mortgage company would profit by $0.015482498847688 - 0.00894034950598 = 0.006542149341708$.

Case 2 (Earlier regime switch from falling to rising house prices): $T_s = 2$, $T = 5$

$$PV(0)_{ABMRS} = 0.276084080324638 \tag{3.5.48}$$

Case 3 (Later regime switch from falling to rising house prices): $T_s = 4$, $T = 5$

$$PV(0)_{ABMRS} = 0.265723482726580 \tag{3.5.49}$$

Case 4 (Longer maturity for the American put option): $T_s = 3$, $T = 10$

$$PV(0)_{ABMRS} = 0.474958294680668 \tag{3.5.50}$$

Notice that when $T = 10$, the corresponding FRM present value at time 0 for the cash flow from 0 to T is 0.486484948666481.

Finally, we plot the expectation of the payment stream a_t with respect to time t in regime switching case (Figure 19) and $P(S(t) \leq M(t))$ (Figure 20).

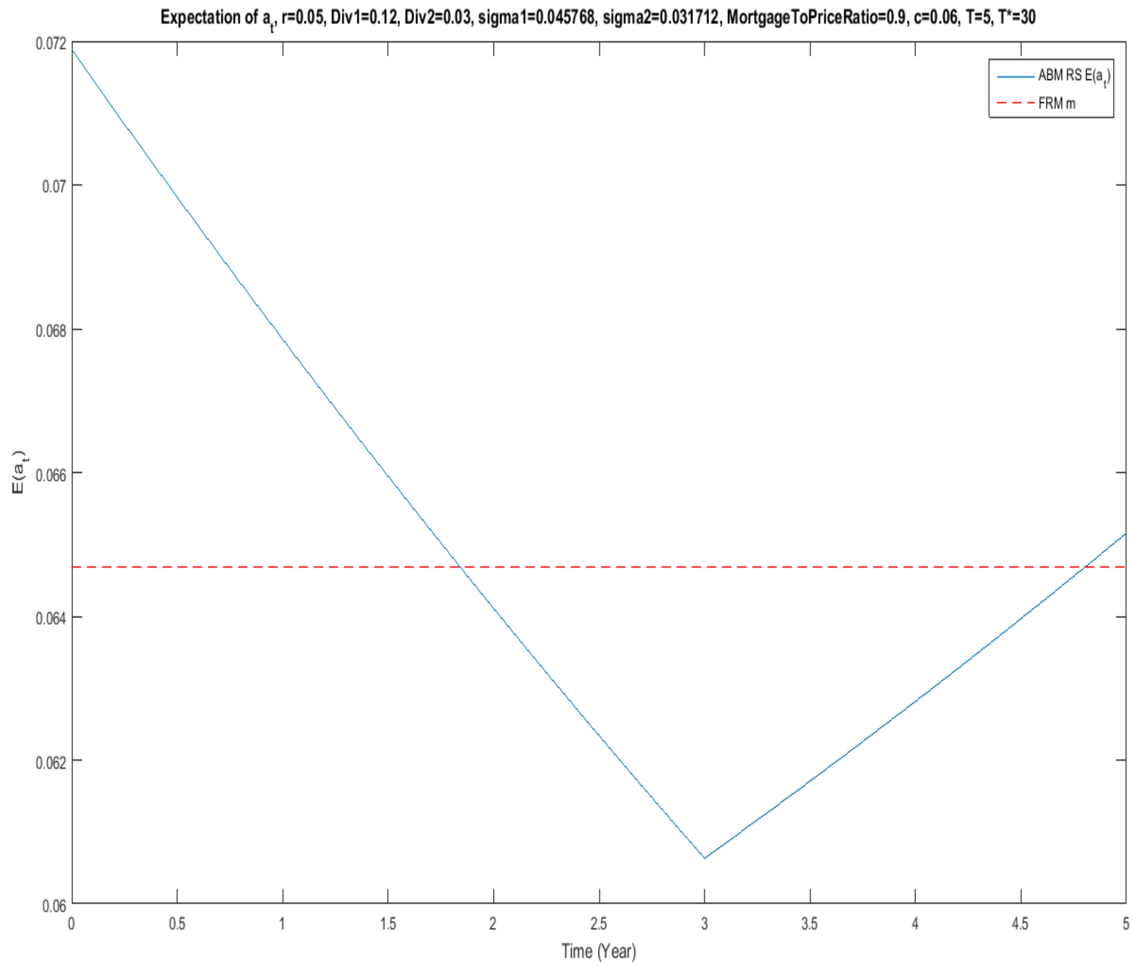

Figure 19: Expectation of a_t, Regime Switching Case. $(S(0) = 1,\ T^* = 30,\ T = 5,\ r = 0.05,\ \sigma_1 = 0.045768,\ \sigma_2 = 0.031712,\ c = 0.06$, Mortgage To Price Ratio $= 0.9)$

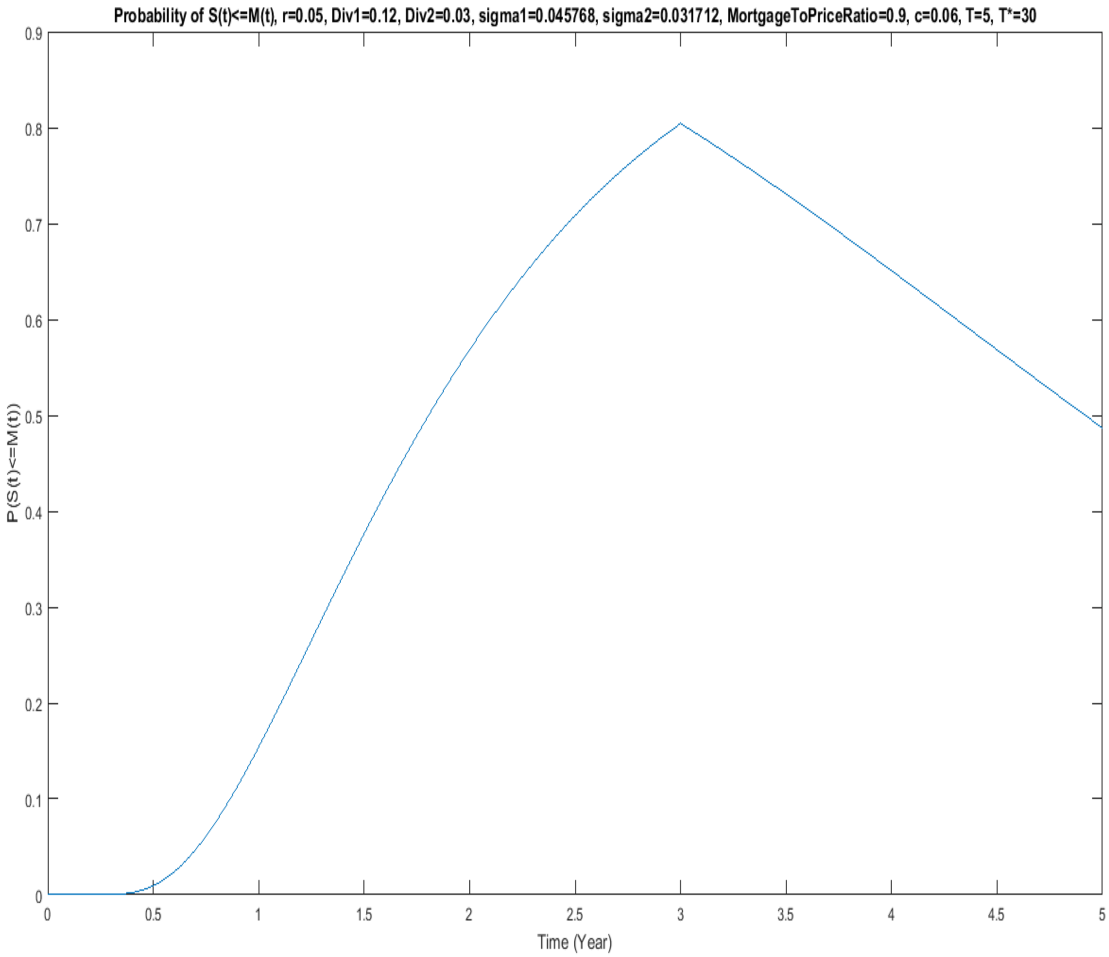

Figure 20: $P(S(t) \leq M(t))$, Regime Switching Case. $(S(0) = 1,\ T^{*} = 30,\ T = 5,\ r = 0.05,\ \sigma_1 = 0.045768,\ \sigma_2 = 0.031712,\ c = 0.06$, Mortgage To Price Ratio $= 0.9)$

Thus we see that there is a high probability that during the end period of declining house prices and the beginning of the housing rebound, the holder of an ABM will see a reduction in mortgage payments of as much as 10 to 15 percent.

4.0 CONCLUSIONS

In the first part of this work we studied the defaults of regional banks due to the housing crisis. Using the publicly available quarterly reports of approximately 6,000 U.S. regional banks, we were able to construct a robust cautionary index that effectively predicted the defaulting banks over 6-month and 1-year time horizons. We provided a PD calculation method for privately-held U.S. regional banks using free and transparent data from the FDIC. We cleaned the data, and applied a bidirectional stepwise logistic regression model to choose the most important predictors at each quarter t. In order to check the predictive power, we apply the estimated coefficients from quarter t to quarter $t+1$. We then assess the goodness-of-fit of the model by three summary measures at each quarter: Area Under the ROC Curve, KS Statistic, and Bad Capture Rate in the Bottom Percentile. The results are shown to be very good for all quarters.

In the second part of this work we studied two strategies for reducing the frequency of home foreclosure (mortgage payment defaults) during the recent housing crisis. In this analysis house prices were modeled by geometric Brownian Motions with time dependent drift and volatility. In an idealized setting, a model foreclosure occurs the first time that the house prices falls below the remaining part of the mortgage. We show that this would occur with very high probability for the data we studied. This would result in very serious credit consequences to many homeowners. To avoid this negative credit impact, the homeowner could purchase a put option allowing him to legally exchange the devalued house for the mortgage with no legal consequences. For the data supplied, we found that this would be a reasonably priced instrument and, more interestingly, that the homeowner would likely not foreclose because (s)he would have had the opportunity to see the housing market begin to rise beforehand. We also studied an alternate strategy, the adjustable balance mortgage, that reduces the mortgage payment when the house price falls below the remaining part of the mortgage. If the reduced payment is calculated on the house price, we found this to be a reasonably priced alternative to encourage homeowners not to foreclose.

5.0 FUTURE WORK

Our work on regional bank defaults suggests that the original objective of identifying a dynamic index for which default arises as a first crossing problem might be unrealistic. On the other hand, some more complicated variant of this structural model setting, similar to those appearing in credit markets [13], might be more appropriate.

Several projects are suggested from the present work on mortgage defaults. One straightforward exercise would be to check the accuracy of our approximating analytic formulas for the probabilities of default by using Monte Carlo simulations [12] with the exact boundaries. Another would be to redo the calculations in the framework when the time of the regime switch, T_s, is random [4]. Finally, it would be interesting to obtain Zillow data on those prices and incidence of mortgage foreclosures from other geographical locations and to assess how well our models fit after homeowner's personal (non-financial) preferences are removed.

APPENDIX

Table 4: Stepwise logistic regression coefficients summary: 1 year default time horizon

Start Quarter t	**beta_0**	**TR**	**EC/A**	**CREL/A**	**CL/A**	**CIL/A**	**GS/A**	**NIM**	**NI/A**
2008.3.31	0.794		-0.449		8.665			-0.643	-118.770
p value	0.485		0.000		0.000			0.005	0.000
2008.6.30	0.334	1.804	-0.538		5.847		-6.216		-19.739
p value	0.778	0.000	0.000		0.000		0.021		0.022
2008.9.30	-2.822	2.009	-0.368		6.309				-44.389
p value	0.000	0.000	0.000		0.000				0.000
2008.12.31	-0.265	0.713	-0.662		6.550				-36.514
p value	0.647	0.002	0.000		0.000				0.000
2009.3.31	4.358		-0.864		7.094		-5.575	-0.623	-66.124
p value	0.000		0.000		0.000		0.003	0.000	0.000
2009.6.30	0.815	0.511	-0.576	2.482	5.527		-4.125	-0.402	-37.086
p value	0.354	0.015	0.000	0.027	0.000		0.029	0.008	0.000
2009.9.30	1.020	0.903	-0.779		3.766				-17.422
p value	0.213	0.000	0.000		0.020				0.013
2009.12.31	0.920	0.734	-0.816				-4.449		-24.276
p value	0.226	0.000	0.000				0.050		0.000
2010.3.31	1.519	0.667	-0.887						
p value	0.047	0.000	0.000						
2010.6.30	0.580	0.685	-0.804						-34.749
p value	0.485	0.000	0.000						0.012
2010.9.30	0.415		-0.875		7.575				-32.031
p value	0.593		0.000		0.002				0.001
2010.12.31	0.316	0.523	-0.927		6.430	6.971			
p value	0.707	0.007	0.000		0.043	0.015			

Table 5: Stepwise logistic regression coefficients summary: 1 year default time horizon (Continued)

Start Quarter t	**beta_0**	**TR**	**EC/A**	**CREL/A**	**CL/A**	**CIL/A**	**GS/A**	**NIM**	**NI/A**
2011.3.31	0.441		-1.000		7.967	8.814	-9.033		-104.500
p value	0.647		0.000		0.034	0.006	0.018		0.001
2011.6.30	4.327		-0.993				-8.321		-34.791
p value	0.000		0.000				0.007		0.025
2011.9.30	1.589		-1.043						-51.385
p value	0.070		0.000						0.000
2011.12.31	1.040	0.367	-1.004						-24.441
p value	0.293	0.024	0.000						0.019
2012.3.31	1.927	0.404	-1.317			9.631			
p value	0.055	0.032	0.000			0.036			
2012.6.30	0.407		-1.027						
p value	0.703		0.000						
2012.9.30	-2.113	0.493	-0.948	5.537					-83.515
p value	0.151	0.031	0.000	0.028					0.000
2012.12.31	-2.342		-0.940	6.296					-47.334
p value	0.116		0.000	0.010					0.002
2013.3.31	-1.890		-1.037			12.748			-96.013
p value	0.252		0.000			0.007			0.019

Table 6: Stepwise logistic regression coefficients summary: 6 months default time horizon

Start Quarter t	**beta_0**	**TR**	**EC/A**	**CREL/A**	**CL/A**	**CIL/A**	**GS/A**	**NIM**	**NI/A**
2008.3.31	1.895		-1.026		8.018		-18.458		
p value	0.178		0.000		0.000		0.048		
2008.6.30	-2.583	1.643	-0.450				-15.162		-31.249
p value	0.121	0.002	0.018				0.028		0.001
2008.9.30	-2.431	1.773	-0.597						-25.873
p value	0.010	0.000	0.000						0.000
2008.12.31	-1.145	0.599	-0.642		4.397				-22.407
p value	0.200	0.024	0.000		0.017				0.000
2009.3.31	1.866		-0.724		5.558			-0.709	-55.145
p value	0.013		0.000		0.001			0.001	0.000
2009.6.30	-0.278		-0.573		3.508	-8.228			-41.131
p value	0.756		0.000		0.046	0.016			0.000
2009.9.30	2.154		-0.966		3.571				-14.037
p value	0.003		0.000		0.041				0.050
2009.12.31	-1.260	0.451	-0.765			5.501			-26.103
p value	0.106	0.002	0.000			0.030			0.000
2010.3.31	2.269	0.570	-1.190				-6.814		-42.666
p value	0.010	0.001	0.000				0.043		0.034
2010.6.30	3.785		-1.470						
p value	0.000		0.000						
2010.9.30	-0.483	0.462	-1.026			10.035			
p value	0.660	0.009	0.000			0.013			
2010.12.31	-0.813	0.889	-1.002			8.604			
p value	0.430	0.000	0.000			0.029			

Table 7: Stepwise logistic regression coefficients summary: 6 months default time horizon (Continued)

Start Quarter t	**beta_0**	**TR**	**EC/A**	**CREL/A**	**CL/A**	**CIL/A**	**GS/A**	**NIM**	**NI/A**
2011.3.31	1.026		-1.525	5.260					
p value	0.325		0.000	0.016					
2011.6.30	1.077		-1.545		13.086				
p value	0.271		0.000		0.010				
2011.9.30	1.166		-1.526						
p value	0.221		0.000						
2011.12.31	0.209		-1.514						-33.309
p value	0.849		0.000						0.009
2012.3.31	3.156		-1.648						
p value	0.008		0.000						
2012.6.30	0.576		-1.389						-80.551
p value	0.575		0.000						0.010
2012.9.30	-1.392		-1.675						
p value	0.559		0.006						
2012.12.31	-0.510		-2.497						
p value	0.775		0.001						
2013.3.31	-0.295		-1.335	6.992					-100.135
p value	0.816		0.000	0.015					0.044
2013.6.30	-2.785		-1.108						-138.504
p value	0.156		0.003						0.001
2013.9.30	1200.300								
p value	0.913								

Table 8: Average monthly house prices in Paradise, Las Vegas

1997-07	109,800	2000-09	130,600	2003-11	171,700	2007-01	285,700	2010-03	126,300
1997-08	110,400	2000-10	131,700	2003-12	176,300	2007-02	284,500	2010-04	125,000
1997-09	110,800	2000-11	132,700	2004-01	180,500	2007-03	281,700	2010-05	123,800
1997-10	111,400	2000-12	133,300	2004-02	184,900	2007-04	277,800	2010-06	122,600
1997-11	112,200	2001-01	134,000	2004-03	192,900	2007-05	273,000	2010-07	121,500
1997-12	113,100	2001-02	135,000	2004-04	204,200	2007-06	269,400	2010-08	120,200
1998-01	113,800	2001-03	135,600	2004-05	216,700	2007-07	268,600	2010-09	119,000
1998-02	114,400	2001-04	135,700	2004-06	229,200	2007-08	266,400	2010-10	118,400
1998-03	114,600	2001-05	136,400	2004-07	238,500	2007-09	260,300	2010-11	117,400
1998-04	114,600	2001-06	137,500	2004-08	244,400	2007-10	252,200	2010-12	115,300
1998-05	114,600	2001-07	138,700	2004-09	248,200	2007-11	244,600	2011-01	112,300
1998-06	114,500	2001-08	139,700	2004-10	250,400	2007-12	242,200	2011-02	110,600
1998-07	114,200	2001-09	140,400	2004-11	251,000	2008-01	239,600	2011-03	110,000
1998-08	114,200	2001-10	140,700	2004-12	253,100	2008-02	232,000	2011-04	109,100
1998-09	114,400	2001-11	141,300	2005-01	257,100	2008-03	224,100	2011-05	107,300
1998-10	114,500	2001-12	141,500	2005-02	261,500	2008-04	221,700	2011-06	106,100
1998-11	114,900	2002-01	141,400	2005-03	264,400	2008-05	217,500	2011-07	105,100
1998-12	115,800	2002-02	141,700	2005-04	266,600	2008-06	209,600	2011-08	103,900
1999-01	117,100	2002-03	142,700	2005-05	267,800	2008-07	201,400	2011-09	102,800
1999-02	117,900	2002-04	143,300	2005-06	268,300	2008-08	195,600	2011-10	102,200
1999-03	118,100	2002-05	143,500	2005-07	269,600	2008-09	191,000	2011-11	101,600
1999-04	118,300	2002-06	143,200	2005-08	272,900	2008-10	184,600	2011-12	100,800
1999-05	118,900	2002-07	143,600	2005-09	278,000	2008-11	177,400	2012-01	100,600
1999-06	119,600	2002-08	144,600	2005-10	284,400	2008-12	169,000	2012-02	101,000
1999-07	120,500	2002-09	146,000	2005-11	N/A	2009-01	161,800	2012-03	101,300
1999-08	121,300	2002-10	147,500	2005-12	289,200	2009-02	155,900	2012-04	101,400
1999-09	122,500	2002-11	149,200	2006-01	290,600	2009-03	150,500	2012-05	102,100
1999-10	123,500	2002-12	150,600	2006-02	292,300	2009-04	144,300	2012-06	103,800
1999-11	123,600	2003-01	151,800	2006-03	295,500	2009-05	140,500	2012-07	105,600
1999-12	124,100	2003-02	152,900	2006-04	296,300	2009-06	136,600	2012-08	107,200
2000-01	125,600	2003-03	153,800	2006-05	294,700	2009-07	132,400	2012-09	109,000
2000-02	126,400	2003-04	154,600	2006-06	293,000	2009-08	129,200	2012-10	111,600
2000-03	126,400	2003-05	155,400	2006-07	291,400	2009-09	127,600	2012-11	114,300
2000-04	126,800	2003-06	156,600	2006-08	289,700	2009-10	126,000	2012-12	116,800
2000-05	127,700	2003-07	158,400	2006-09	288,600	2009-11	125,300	2013-01	118,800
2000-06	128,600	2003-08	160,700	2006-10	287,700	2009-12	125,100	2013-02	121,600
2000-07	128,900	2003-09	163,700	2006-11	286,100	2010-01	125,600	2013-03	124,600
2000-08	129,500	2003-10	167,600	2006-12	285,600	2010-02	126,400		

BIBLIOGRAPHY

[1] E. I. Altman. *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting and Avoiding Distress and Profiting from Bankruptcy*. New York: Wiley, 1993

[2] B. W. Ambrose, R.J. Buttimer Jr.. The Adjustable Balance Mortgage: Reducing the Value of the Put. *Real Estate Economics*, Volume 40, Issue 3, pages 536-565, Fall 2012

[3] J. F. Bovenzi, J. A. Marino, and F.E. McFadden. Commercial bank failure prediction models, Economic Review. *Federal Reserve Bank of Atlanta November* 1983: 14-26.

[4] J. Buffington and R. Elliott, American Options with Regime Switching. *International Journal of Theoretical and Applied Finance*, 2002; 5(5): 497-514.

[5] X. Chen, H. Cheng, and J. Chadam, Non-Convexity of the Optimal Exercise Boundary for an American Put Option on a Dividend-Paying Asset. *Mathematical Finance*, Volume 23, Issue 1, pages 169-185, January 2013

[6] P. Ciurlia, A. Gheno, A Model For Pricing Real Estate Derivatives With Stochastic Interest Rates. *Mathematical and Computer Modelling*, Volume 50, Issues 1-2, Pages 233-247, July 2009.

[7] P. Carr, R. Jarrow, and R. Myneni. Alternative Characterizations of American Put Options. *Mathematical Finance*, Volume 2, Issue 2, pages 87-106, April 1992.

[8] K. E. Case, and R. J. Shiller. Prices of Single Family Homes Since 1970: New Indexes for Four Cities, *New England Economic Review*, 45-56. 1987.

[9] P. Gelain, and K. J. Lansing, House Prices, Expectations, and Time-Varying Fundamentals. *Journal of Empirical Finance*, Volume 29, Pages 3-25, December 2014.

[10] R. A. Gilbert, A. P. Meyer, and M. D. Vaughan. The role of Supervisory Screens and Econometric Models in Off-site Surveillance. *Federal Reserve Bank of St. Louis*, November/December 1999.

[11] R. A. Gilbert, A. P. Meyer, and M. D. Vaughan. The role of a Camel Downgrade Model in Bank Surveillance. *The Federal Reserve Bank of St. Louis*, August 2000.

[12] E. Gobet, Advanced Monte Carlo Methods for Barrier and Related Exotic Options. *Mathematical Modeling And Numerical Methods In Finance*. January 2008.

[13] J. Huang, Hierarchical Structure Models of Portfolio Credit Risk. PhD Dissertation, University of Pittsburgh, 2013.

[14] L. Jiang. *Mathematical Modeling and Methods of Option Pricing*. World Scientific Publishing Company. Jul 2005.

[15] Chiong-Long Kuo. Housing Price Dynamics and the Valuation of Mortgage Default Options. *Journal of Housing Economics*, Volume 5, Issue 1, Pages 18-40, March 1996.

[16] J. Krainer, S. F. LeRoy, and M. O, Mortgage Default and Mortgage Valuation. Federal Reserve Bank of San Francisco Working Paper. November 2009.

[17] A. E. Kocagil, A. Reyngold, R. M. Stein, and E. Ibarra. Moody's RiskCalc Model For Privately-Held U.S. Banks. Moody's Investors Service July 2002.

[18] R. C. Merton. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29 (1974): 449-470.

[19] P. J. Schonbucher. *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley. June 2003.

[20] S. Shreve. *Stochastic Calculus for Finance II, Continuous-Time Models*. Springer-Verlag New York. 2004.

[21] P. Willen, C. Foote, K. S. Gerardi, and L. Goette. Reducing Foreclosures. Federal Reserve Bank of Boston Public Policy Discussion Papers Series, paper no. 09-2 (2009).