Mapping the Drift to Default:
A credit risk modelling approach to the early termination of UK residential mortgages.

Steven Frank Kay

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Management Science and Business Economics

The University of Edinburgh.

2013.

“*Applied instruction is excellent training for those who are resolved of their own accord to make up the missing links in their education by private study*”

Albert Kesselring. The Memoirs of Fieldmarshal Kesselring, 1963

DEDICATION

To my Wife Irene, always there with help and care, a smile for all.

To my late parents, who taught me to use my head and . . .   . . . to follow my heart.

On the seas and far away,
On stormy seas and far away;
Nightly dreams and thoughts by day
Are aye with him that's far away.

From On the Seas and Far Away ROBERT BURNS (1794)

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

I share this accomplishment with many.

My first and foremost acknowledgement must go to my advisor and mentor, Prof Jonathan Crook. I can honestly say that what I have accomplished would not have been possible without his guidance and always generous support. I would also like to thank my other advisor and mentor, Prof Jake Ansell for his attention to detail and thoughtful participation in this process. I am grateful also to the assistance of the financial institution, whose wish is to remain anonymous and who provided both encouragement and the data for this thesis

I would like to include a special thanks to Ross Gayler and Dr. Tony Bellotti for their good counsel that has served me well over my years in the doctoral program and to the other members of the Credit Research Centre, University of Edinburgh and associated research group for their active encouragement and always stimulating exchanges.

I also acknowledge the generous help of Damien Butters for his honest and useful critique of the mathematics in my earlier activities towards this thesis.

There were many, many, others, outside those most closely and academically connected to my thesis, that also provided important support during this process. Thanks to all those with whom I have shared my graduate school experience, they have been a source of comfort and cheer, ensured I took a break from study and in particular provided the odd sanity check throughout my labours.

Finally, I would like to give grateful thanks to the EPSRC and Credit Research Centre, University of Edinburgh for their very generous funding, without which this work would not have been completed.

Declaration of Own Work

I can confirm that no portion of the work referred to in this thesis has been submitted by me in support of an application for another degree or qualification of this or any other university or other institution of learning in the UK.

Steven Frank Kay

April 2013

ABSTRACT OF THE THESIS

Mapping the Drift to Default:
A Credit Risk Modelling Approach to the Early Termination of UK Residential Mortgages.

By

Steven Frank Kay

Doctor of Philosophy in Management Science
University of Edinburgh, Edinburgh, 2013

This thesis is devoted to UK Mortgage Performance Modelling.

The research conducted uses an option pricing methodology to model theoretically the value of Mortgages, the Option to default and the probability to default and to compare the predictive accuracy of the latter with the predictive accuracy of data driven credit-scoring techniques.

Theoretical models are constructed to represent the life cycles of loans collateralised by real property operating within a stochastic economic environment of house-price and interest rate. These realistic mortgage models provide a confirmation of recent research based upon a relaxation of the assumption of financially rational, 'ruthless' prepayment, bridge a potential oversight in existing research by an extension of existing modelling in the stochastic behaviour of the house price process and present a proposal for a straightforward approach utilising characteristic measures of borrower delinquency and insolvency that enables estimation of the default probabilities implicit in residential mortgages using a simple but enhanced optimising structural model.

This model straightforwardly demonstrates that one can predict the probability of eventual default, beginning at the origination of the loan, the time when a lender would be most interested in making such a determination.

Secondly the problem of mortgage loan default risk is empirically assessed in a number of different ways focusing upon analysis of the competing risks of early termination, the inclusion of macro-economic variables - time varying covariates and unobserved borrower heterogeneity.

Key insight is provided by means of a multi-period model exploiting the potential of the survival analysis approach when both loan survival times and the various regressors are measured at discrete points in time. The discrete-time hazard model is used as an empirical framework for analysing the deterioration process leading to loan default and as a tool for prediction of the same event. Results show that the prediction accuracy of the duration model is better than that provided by a single period logistic model. The predictive power of the discrete time survival analysis is enhanced when it is extended to allow for unobserved individual heterogeneity (frailty).

Chapter 1

Introduction

The UK residential mortgage market has received a great deal of attention in the past decades due to its importance to the economy. Firstly, the residential mortgage market stimulates the economy by creating jobs, wages and tax revenue through the construction of new homes. Secondly, the residential mortgage market strengthens social stability by increasing and expanding homeownerships across the UK. Lastly, the residential mortgage market is crucial to the prosperity of the British economy.

Now even amid all the changes and controversies prevalent in the financial sector at this time, lenders continue to originate mortgage loans because the interest received as payment by the borrower over the lifetime of the loan is profitable.

A simple numerical illustration of such profitability is given as Appendix A. In addition to such profitability the loan portfolio at a large banking institution contains a high component of residential mortgages making the issue of appropriate capital for residential mortgage credit risk important.

The importance of the real estate and mortgage markets is thus widely recognised in well developed economies. It is primarily for these reasons that, research into methods for rational pricing of mortgages and mortgage backed securities and the evaluation and modelling of the risk attached to mortgage contracts is being undertaken by many of the larger banks and building societies.

There are two major sections in this thesis. First, theoretical models of mortgage value are presented and a new extension to this theory is described. The second sheds light on the plausibility of such models by empirically parameterising models of default and prepayment

1.1 Background and Research Context

To give some appreciation of the magnitude of the outstanding debt on residential mortgages, according to the Council of Mortgage Lenders (CML), in the year ending September 2011 gross mortgage lending reached an estimated £148.2bn approximately and the total value of the mortgage loans market at £1.2 trillion owed by British borrowers, is representative of 11.3 million mortgages.

In terms of a credit risk assessment the Council of Mortgage Lenders estimates for 2011 was that 1.58% of all mortgages were in arrears and 0.35% of all mortgages were in repossession. Repossessions are invariably the consequence of an extended period of household financial distress during which time borrowers have accumulated arrears, the level of which results in the default event.

Bearing in mind the foregoing this thesis is devoted to Mortgage Performance Modelling. It opens with a review and critical evaluation of issues that are considered to be relevant in a survey of both Theoretical and Empirical work on the rational pricing of Mortgages and Mortgage Backed Securities as derivative assets[1], referred to as the option-pricing approach to mortgage valuation.

---

[1] A Derivative Asset is one that is not necessary to describe the underlying real economy. It is considered to be a redundant asset, with its value dependent, entirely on the variables that do influence the state of the economy. This basic assumption drives the option theoretic approach to mortgages, in that, mortgages depend upon the real economy through the term sructure of interest rates and the dynamics of house prices but they are not necessary to determine this underlying economy.

In this approach the models that are constructed are used to represent life cycles of loans collateralised by real property operating within a stochastic economic environment.

The research accent will be on an examination of the nature of the determinants that underlie the possibility of terminating such mortgages by either prepayment or default with primary attention being given to default. Within this framework an analysis of the strengths and weaknesses of the option-pricing approach to mortgage valuation will be given.

A further motivation behind this research is a need to better understand option-pricing techniques for the valuation of mortgages. For, more importantly, it is the solution to the valuation problem that is a pre-requisite to obtaining the Probability of Default (PD) and hence the Loss Given Default (LGD) and it is especially important to consider this under the Basel II Banking Supervision rules.[2]

1.1.1 Mortgages: Overview of the Contracts and the Markets

This section presents the fundamentals of mortgage contracts.

A mortgage is a financial contract that falls under the umbrella of a fixed income product. It is a contractual agreement between a lender and a borrower, where the lender loans the borrower the capital needed to purchase a property (which may be residential or commercial). The principal of the loan is secured against the specified property and is paid back in instalments (these are usually monthly) whose

---

2 The Basel Committee on Banking Supervision is responsible for setting current capital standards for internationally active banks regarding the amount of capital that must be put aside to deal with credit risk.

amortised value includes the total loan amount (that is, the principal) plus interest due.

The Residential Mortgage is one of the most common and oldest of financial contracts. Despite its popularity, it presents some features that are relatively unusual in financial markets and these make it especially difficult to evaluate. Residential Mortgages are not *securities* since they are not (usually) traded. However, this type of debt instrument can be treated as a derivative security - it may be viewed as a fixed income instrument combined with the embedded options to prepay and to default.

A basic principle of financial investment is that the value of an asset is the present value of its future cash flows. In the case of Mortgages and Mortgage Backed Securities, accurate estimation of loan prepayment and default is essential to investors in their forecasting of expected mortgage cash flow patterns. Couple this with the desire to gain an accurate understanding of the risks involved in prepayment and default as represented by loan and borrower characteristics and you have formed the very heart of the residential mortgage termination research literature.

Before describing the financial modelling undertaken in Chapters 3-5 it may be helpful to quickly review the most relevant elements in the valuation of a financial contract.

The valuation of any financial asset depends upon two main factors

- Its cash flow structure and
- The economic environment in which it is traded

The contract rate and the other contractual features of the loan determine the cash flow structure of a mortgage. Some important contractual features are characteristic of almost all types of mortgages, of special relevance is the fact that the term to maturity is particularly long - reaching usually 25 years in the UK and 30 years in the USA. Other very important features generally encompassing all residential mortgage contracts are the possible sources of early termination: The borrower is given the option to terminate the loan prior to the maturity date, either through pre-payment or through default. Very many other contractual features vary with the specific characteristics of the underlying contract and require individual specifications. There exist nowadays several different types of mortgage contract with the basic classification being based upon how the loan is repaid:

Repayment: each monthly instalment pays off an amount of the capital plus the interest accrued. During the early years of the contract almost the entire amount paid consists of interest, but gradually as the unpaid balance decreases, the proportion of the payment consisting of interest decreases, until at the final payment, no more is owed. Since repayment type mortgages do not have any in-built life assurance cover it is relatively common practice for the holders of this type of loan to take out a separate life insurance policy that covers against death.

Interest only: each monthly instalment just pays off the accrued interest on the amount borrowed with no repayment of the capital sum (i.e. a non-amortising loan) and ideally an amount of money is put into some other separate investment vehicle which should be able to pay the loan capital at the maturity of the mortgage contract (i.e. the endowment policy premium) The lender is the beneficiary of the endowment policy.

If the borrower dies, the issuer of the insurance policy pays off the debt (life insurance component of the endowment policy). If this does not happen during the term of the loan, and the borrower also does not default, the insurance policy accumulates the necessary value to pay off the loan at the maturity leaving the borrower owning an unmortgaged house (investment component of the endowment policy). This may involve a degree of risk in that, at maturity, the value of the investment component may be insufficient to repay the full amount of the debt.

The main element that exerts a significant influence over the cash flow structure is the type of contract interest rate inherent to the mortgage. In a contingent claims valuation framework, the elements that are used to characterise the economic environment are the most relevant sources of risk inherent in the asset under study.

Turning now to the UK mortgage market, which has three segments to its structure as outlined below:

i. Prime: Generally serving borrowers with an unblemished credit history. These are standard (underwritten criteria) mortgage loans with no high risk features and the normal payment characteristics required of a fully amortising loan.

ii. Alt-A: Generally serving borrowers with credit histories close to prime but the loans have one or two more high-risk features such as a limited documentation of income or assets

iii. Sub-Prime: Generally serving borrowers with blemished or limited credit histories and these loans have associated higher-risk features, typically higher interest rates and higher charge fees in comparison to prime loans.

There are two major groups of providers on the supply side of the UK residential mortgage market: building societies and banks. Historically building societies have dominated the market, but with liberalisation of the financial services sector (this took place in the 1980's) and the transformation of some of the biggest building societies into banks, this general picture has changed significantly over the last decade or two.

Building societies tend to set their mortgage rate relative to the evolution of the prevailing money market rates and banks determine their mortgage rates as a function of their respective base lending rates which are indirectly related to the prevailing money market rates. Thus regardless of the type of lender, the evolution of the prevailing money market exerts an important influence on the specification of the mortgage contract interest rate. Interest rates for mortgage contracts can be either:

Fixed: Monthly payments are independent of the evolution of the money market interest rate for a pre-determined period of time. The rate is fixed for a set period of time, 2-5 years initially, and during this period of time, mortgage payments only depend upon the rate that was agreed between the parties in the contract at origination. Fixed Rate Mortgages are the dominant type of collateral in the mortgage securitisation market.

Variable: Monthly payments made by the borrower evolve in accordance with the lenders mortgage rate. The interest rate is set at between 1% and 2 % above the bank base rate and in the UK the most common rate adjustment regime allows for the contract rate to be changed whenever the lenders mortgage rate changes. These mortgage loans tend to have an initial period of fixed interest rate usually two or three years but then 'revert' to a variable rate for the remainder of the term of the loan.

Lending institutions are usually specialised in the management of interest rate exposure. A numerical example illustrating the basics of mortgage contracts is given in Appendix B.

1.1.2 A brief summary of recent events in the mortgage markets

The next few paragraphs complete this section by briefly considering the conditions in the UK mortgage market during the study period and more recently (that is from 2000 up-to the first half of 2012) with links to reasons for the current credit crisis.

To give some appreciation of the magnitude of the outstanding debt on residential mortgages, according to the Council of Mortgage Lenders (CML), gross mortgage lending reached an estimated £148.2bn in the year ending September 2011. In September alone this figure stood at £12.9bn, down 18.9% on £13.15bn in August 2011, but 3.7% up from September 2010, when it stood at £12.44bn.

The CML also revealed that gross mortgage lending amounted to £38.6bn in the third quarter of 2011 compared to £33.5bn in the second quarter of 2011.

There appears to be no clear trend from one month to the next in terms of mortgage lending.

Mortgages are a type of credit, so they were adversely affected by the 'credit crunch', that occurred pre-recession, in 2008. During this period, large amounts of credit were withdrawn rapidly by banks and other lenders, as uncertainty grew over the abilities of borrowers who already had mortgages to pay back their debts. This led to a sub-prime mortgage lending crisis, where borrowers defaulted on loans as they were unable to keep up their mortgage repayments. Such loans were set up by banks when credit was much easier to come by and the conditions of loans were less stringent.

The overall market is however far from consistent, with the number of properties repossessed standing at 9,200 in the third quarter of 2011, up from 9,100 in the second quarter. Despite this, the total number of mortgages in arrears fell, between the second and third quarters of 2011, with the total number of mortgages with arrears of 2.5% or more of the outstanding balance falling to 161,600 (1.44% of all loans), down 2% from 165,200 (1.47% of all loans) over the same period. For some perspective it is noted that despite the depth of the recession in 2008, the mortgage default rate in the UK has been lower in the current recession than that in the early 1990s. The peak for home repossessions occurred in 1991, when over 70,000 homes were repossessed or 0.77% of outstanding mortgages. The peak in the current recession was 46,000 home repossessions in 2009. (0.4%)

The real problem behind the credit crunch is the repossessions in the US.

It was the defaults on the sub-prime mortgage loans that have caused a £1.3trillion black hole in the global financial system. The UK is suffering not because UK repossessions are high, but because British banks were exposed to the credit defaults in US banks.

What were main causes of Default in 2008-12?

- Interest rates

One factor that clearly explains different mortgage default rates is the base interest rate. In the early 1990s UK interest rates were very high, reaching a peak of 15%. The high interest rates were one of the main causes of the then recession. High interest rates caused tremendous problems for homeowners. The sharp rise in interest rates saw the cost of mortgages almost double. This meant many borrowers could not afford the repayments and so they went into default.

The current recession was caused by the credit crunch (not high interest rates). In fact the Bank of England cut interest rates to 0.5% (an historic low) to try and stimulate the economy. This favoured many borrowers particularly those on tracker and variable mortgages who saw their monthly mortgage payments fall. Therefore, although there was a rise in unemployment and low income growth, at least interest rates were low.

One factor about interest rates is that not every mortgage holder saw their mortgage rate fall by as much as base rates. Many banks took the opportunity of keeping mortgage rates high and not passing the base rate cut on to consumers.

Therefore, mortgage rates were still high. But, quite significantly lower than in 1991.

- Unemployment.

In this period, UK unemployment rose significantly. Unemployment is the biggest cause of mortgage defaults as people receiving job seekers allowance see a fall in disposable income and may no longer be able to afford repayments.

- Falling real incomes

The 2008-10 economic slowdown was a period of stagnant or falling real incomes - value of wages adjusted for inflation. For one of the first times, people in work saw negative real wage growth. Inflation was higher than nominal wage growth leading to squeeze on living standards. This squeeze in real incomes makes it more difficult for people to meet mortgage repayments.

- Size of Mortgages

During the boom years of 2001-2007, banks had lent unorthodox mortgages such as four or five times income multiples and self-certification mortgages. This meant many homeowners were paying a large percentage of their income on their very high mortgages. From 2004 the increasing debt service ratio and increasing proportion of properties in negative equity were key drivers of the increase in mortgage possessions.

- To an Optimistic Future and Beyond

In 2012, the Council of Mortgage Lenders (CML) has predicted that there will be 900,000 residential property transactions, up from 860,000 in 2011. After this period, a slow recovery for the mortgage market is to be expected thanks to continued economic uncertainty.

However, by the end of 2014 it is estimated that the market will start to recover faster; this mostly due to the stricter regulations now governing mortgage lending — especially subprime loans — in the wake of the financial crisis.

Summarising, the mortgage crisis that has shaken the financial stability of many developed countries in 2007 and 2008 has highlighted just how important it is to accurately assess the risks in mortgage lending, in order to price these risks correctly.

1.1.3 Mortgages and Option Pricing

With fluctuations in the volume of mortgage trade, more especially with recent turbulent years in both the primary and secondary markets, there is the need to rationally price mortgage assets and formulate assessments of the risks involved in holding a portfolio of these assets. The markets treat mortgages as financial derivative assets, and quite naturally they employ the use of option pricing approaches in the valuation of these derivative assets.

The theoretical foundations of this research focus quite specifically on the use of option pricing methodology for the valuation of mortgage contracts and this section reviews the option pricing method.

Briefly, any right or set of rights that gives its holder the possibility of making future decisions that are capable of affecting the value or the timing of the cash-flows from a venture has value. The Contingent Claims approach to the valuation of financial assets takes into account this fundamental principle, and the notion of arbitrage free markets, in order to produce a theoretical framework under which it is possible to value assets with implicit or explicit option characteristics.

For the options approach to be valid it must be able to predict if a mortgage holder will terminate early their contract by either prepaying (An American Call type option) or defaulting (A European Put type option). A brief introduction to the basic ideas of these financial options and terminology associated with them is given as Appendix C.

To give due consideration to the decision of early termination, an understanding of borrower behaviour is crucial. One can imagine many possible personal ‘non-financial’ reasons why a mortgage holder may so easily prepay (or sell) the house which makes it less plausible that default will ever occur unless the mortgage becomes more costly than the underlying property. The options approach has the ability to endogenously predict such ‘rational’ default behaviour.

1.1.4 Research Approach and Methodology

As mentioned earlier, this thesis presents new contributions to the option pricing theory of mortgage default and secondly new empirical models that predict time to mortgage default.

In very general philosophical terms the research paradigm in which this thesis sits neatly is termed ‘*Positivism*’ and is based exclusively upon quantitative research practices generally made using scientific methods. (For the interested reader a fuller description can be found in Chapter 4, pp 84-113 of Saunders, et al, 2000).

In the last 20 - 30 years the theoretical pricing models of mortgages as derivative assets has become widely accepted by the financial community. Now, whilst the application of financial mathematical methods to price options is relatively recent (dating back to the 1970's), the fundamentals of option-based pricing models can be traced back to a far earlier origin in economic research.

It was the French mathematician Louis Bachelier at the turn of the twentieth century who pioneered the *Random Walk*, as a mechanism for the description of the movements in stock market prices. This innovation actually predates the famous work of Albert Einstein who laid the scientific foundations on *Brownian motion* named after Robert Brown, a Scottish botanist, who observed the random motion of small particles suspended in a fluid – and it is this that is the idea of a Random Walk for physical processes.

The starting point for almost all mortgage valuation models is a focus on the underlying dynamics of the movement in Interest Rates and Asset prices. The processes that model the uncertainty in these include a *Wiener process* as the representation of the 'randomness' aspect involved. Now in relation to mortgages, two forms of uncertainty are present:

- <u>Term Structure Risk</u> - As interest rates are not a directly traded asset, an *Equilibrium Model* can be used to value interest dependent contracts, which means that neither party to the mortgage would enter into the contract unless it was fair at the start. In other words the value of the mortgage to the lender at origination must be equal to the loan principal lent to the borrower. The term structure is typically captured by a single variable, the Spot Rate, which results in a single market price of risk. The *Local Expectations Hypothesis* assures us that this market price of risk disappears.

- <u>Default Risk</u> - House price is the source of the credit risk in default. The house asset is itself a tradable asset (albeit not in a standardised form and in a relatively illiquid market (infrequently traded), the analogy with the option on a dividend-paying stock is quite close.

In the 1950's, the foundation for option pricing theory was laid, that is an approach to derivative pricing via *expectations, real as opposed to risk-neutralised ones* was considered . Also during this time, the relationship between a stochastic differential equation for some independent variable and the stochastic differential equation for a function of that variable was identified.

In the more recent times work on option pricing theory produced closed form solutions for the valuation of a European Call Option on an underlying asset, in which the interest rate was considered constant. Following on from these seminal contributions, a whole new approach to finance has been under development, both in extending and applying this conceptual framework to the valuation of many different types of explicit and embedded options.

Unfortunately, solutions to the valuation problem posed by the complex contracts that form the basis of mortgages do not exist in a closed form and to value these, one must generally resort to one of the plethora of numerical solution techniques. These include Monte Carlo Simulation methods and *Finite Difference Techniques* from which good approximations of mortgage values can be obtained.

The work presented in this thesis is thus both well informed and well guided by this whole spectrum of established mathematical theory and practical option pricing techniques. The work thus sits firmly in a coherent and useful framework in which the mortgage contract is treated as a derivative asset. The research approach adopted in this thesis is two-fold and is comprised of an initial mixed method analytical mathematical model development followed by supportive empirical data driven statistical modelling methods. Specifically the research encompasses an element of theoretical modelling and its practical application and an element of practical data analysis, both elements are briefly signposted below.

The theoretical modelling makes use of a general valuation framework based on the Cox Ingersoll and Ross, (1985a), (CIR) equilibrium model – details of which are provided elsewhere in the thesis. The valuation equation so derived provides the value of the assets under study given certain terminal and boundary conditions. The intrinsic complexity of mortgage contracts with embedded options to default and prepay leads to the situation where no closed form solution to the problem can be obtained. Consequently it is necessary to utilise a numerical solution technique, based upon the finite difference method. Using this methodology a series of simulation exercises are performed, the results generated identify the sensitivity of the different assets to changes in the various parameters of the model. These results are found to be entirely consistent with economic reasoning. A number of adaptations to the model follow providing further confirmation of recent research in the field of application.

1.1.5 Aims, Objectives and Thesis Structure: Structural Modelling

Chapters 3-6 present a realistic option theoretic mortgage valuation model based on stochastic modelling. The main objectives of this are as follows:

a. To use a competing risks contingent claims framework for valuing the mortgage and the associated embedded options based upon the two standard state variables, namely interest rate and house price and including the potential for early termination and the valuation of associated insurance products.

b. To verify the importance of financial option variables in mortgage terminations and demonstrate the effects of transactions costs of early termination upon a mortgage contract's value.

Chapter 7 aims to address several perceived limitations of this benchmark model. The main objectives of this chapter are

a. to bridge a gap between theoretic and empirical models by investigating the use of a stochastic house price process that is mean reverting.

b. to verify recent research findings by considering the use of a barrier type option as a replacement for the American Call option in developing new valuation approaches for mortgage products with uncertain time of termination through prepayment.

c. to appeal to a mechanism that will allow the effects of macroeconomic variables to be potentially investigated by the incorporation of a *discount factor* that will, in general, decrease the value of the mortgage depending on the most current economic conditions available.

In light of recent turbulence in the financial markets the thesis also aims to determine the role of changing house prices, the main objective being:

d. to suggest a framework that links causes and effects by utilising the traditional 3C's of credit underwriting, expressed through an affordability indicator and the loan to value ratio as a determinant of credit risk, coupled with an illustrative application to real UK market data.

In general the thesis discusses and takes the reader through the whole model building process, discussing the performance of the models where this is feasible and considering selected risk and portfolio optimisation topics. It therefore aims to illustrate the development of a complete and consistent pricing model for different contract specifications with the main objective of identifying how its fine-tuning to the pricing of MBS may be achieved. Concerning MBS, the primary objective is to show how to improve on existing pricing models with respect to the challenges associated with MBS valuation

1.1.6 Aims, Objectives and Thesis Structure: Data Driven Modelling

The empirical research of this thesis investigates potential predictive factors for survival in mortgage loans through the construction and validation of statistical models evaluating contractual, financial option related and borrower behaviour related factors, whilst investigating the underlying assumptions, appropriateness and suitability of such models thus supporting the structural modelling of previous chapters.

Chapters 8 and 9 aim to address, in a systematic way, the structure of influences affecting default and prepayment risks as evidenced in real financial data supplied by a mortgage loan originating market participant. It is a robust dataset containing a reasonable event rate and a relatively long follow-up of active loans.

The main objectives of this section are as follows:

a. To capture the paths of economic variables and appropriately link them to borrower early termination decisions through the use of Time Varying Covariates in the context of a Discrete Time Survival Analysis that links the probability distribution of a duration variable to a set of such covariates.

b. To investigate the early termination decision choices of mortgage loan borrowers in the context of a Mixed Logit Model that includes both the characteristics of the choice alternatives – the Competing Risks nature of default and prepayment - and characteristics of the individual.

c. To provide a more qualitative prepayment and default risk assessment by using an on-the-fly Cumulative Incidence investigation of the survival data.

d. To extend the competing risks survival modelling approach to allow for unobserved borrower heterogeneity and correlation over time - since a parsimonious set of covariates cannot completely explain all the variability in observed time to default and mortgage loans with a similar grouping structure may have dependent survival times.

In summary, the purposes of the reduced form modelling is to examine the determinants of mortgage failure from the pool of financial, borrower and market driven variables, and to explore whether incorporating macroeconomic patterns improve forecasting results.

1.1.7 Overview of Empirical Methodology

In this thesis what is meant by *credit risk* is the following definition ...

'Credit Risk is the risk of loss due to a borrower's non-payment of their Mortgage loan or other associated lines of credit'

derived from a more general expression of Risk under Pillar 1 of the Basel II Capital Accord, (BIS, 2001)

Central to credit risk is the default event, which occurs if the debtor is unable to meet his/her legal obligation according to the debt contract's terms and conditions. An example of the default event of interest here is mortgage foreclosure. Other forms of credit risk include: repayment delinquency in mortgage loans and the severity of loss upon the default event. An abundance of literature fostered by both academics in finance and practitioners in industry inform the data analyses undertaken in this research. Now there are two parallel dimensions in this universe of knowledge based upon a dichotomy of data availability:

a) Direct measurement of credit performance leading to a stream of activity in modelling '*actual*' default probability corresponding to the direct observations of defaults (physical default in finance terms) the other

b) Based upon prices observed in credit markets leading to a stream of activity in modelling the '*implied*' risk-neutral probability of default from the credit market data.

In this thesis an attempt is made to develop statistical models based upon actual credit performance data. Now loosely speaking default risk models based upon actual credit performance data belong exclusively to survival analysis, since by definition it is the analysis of time-to-failure data. As an approach to credit risk modelling survival models offer important advantages

There is flexibility

i) in parameterisation of the default intensity – lifetime distributions for the Hazard Rate include: Exponential, Weibull, Lognormal, Log-logistic, and Generalised Gamma all of which are supported by the SAS statistical package

ii) in incorporating various types of covariates – Discrete-time, Continuous-time, Time independent, and continuous or dichotomous.

There is also effectiveness in modelling credit portfolios, they are straight forward enough to implement and use and so may enrich or extend a family of credit risk models.

With reference to the three empirical themes identified above:

- Focussing on Time-Varying Covariates:

The benefits of modelling with the Cox Proportional Hazard Rate model include the ability to associate loan characteristics with prepayment and default rates and an easy adaption of the model to include time-varying macroeconomic covariates.

Now since it known that the Cox model can be written equivalently as a logit model with a logistic transformation of the hazard rate, (instead of the odds ratio) and dependent upon convergence of the two specifications as a the hazard becomes increasingly small. A logit model was also developed bearing in mind that in mortgage analysis the hazard rate is not necessarily small – since prepayment is for instance a relatively frequent event compared to default which is a relatively rare event in the supplied dataset.

- Focussing upon the Competing Risks of Default and Prepayment:

In this theme the typical '*cause-specific*' approach for competing risks is initially performed. A survival analysis is undertaken for each mortgage termination event type separately, with the other (competing) event type being treated as censored. In this basic hazard rate model the competing risks are estimated independently. However, in a contingent claims setting the borrower is assumed to hold a *joint* termination option encompassing both default and prepayment, indicative of the situation that the independence assumption is not appropriate. To make allowance for this, two alternative approaches are considered. The first is to make use of the marginal probability of survival via calculation of Cumulative Incidence. This is particularly appropriate as the competing risks are mutually exclusive and the termination events are non-recurrent, that is, in the data one and only one of the events occurs at any one time and only once over time. The second is to consider a multinomial logit model which encompasses an expanded set of mortgage decisions -to include delinquency- allows for conditioning upon past decisions and allows for correlation among the competing risks.

- Focussing upon Unobserved Heterogeneity and Correlation over Time.

In this theme unobserved heterogeneity and correlation over time - frailty - as a representation of an individual mortgage loans' vulnerability to the risk of default or prepayment, captures the total effect of all those influences that impact on the individual mortgage loan's risk of termination that are not included in the baseline hazard function or the information set conveyed by the observed covariates. Such random effects are not observation specific, but instead are '*shared*' across groups of observations, causing those observations within the same group to be correlated.

The Standard Cox PHM analysis assuming independence of the survival times is representative of the '*No Frailty'* case and a Stratified Cox Proportional Hazards Model (SPHM) is used to control for the within-group correlation serving as an extension of the standard Cox's (PHM) model. In this '*Shared Frailty'* model, to examine whether mortgages are associated in their survival times, the mortgages were grouped by UK Region consisting of a county or a group of counties where they were originated. This stratification allows for the control of spatial variation for the local economic risk and unmeasured socioeconomic factors. Here frailty may be considered as the combined effect of shared demographic and regional policy effects on individual borrowers. In addition the survival model with frailty is used to investigate the potential correlation of mortgage defaults for mortgages originated in a given year and stratify the mortgage sample by the borrowers' initial LTV ratio.

1.2 Contributions to Knowledge

This thesis makes the following contributions to knowledge

1. We confirm theoretical solutions and relationships concerning mortgage valuation and mortgage default.

2. We present an option pricing model of mortgage valuation based upon house price following a mean reverting process and show that certain elements of local house prices may be more important for mortgage valuation than previously thought.

3. We present a new method to incorporate observable economic conditions into the theoretical options models of mortgage value and default.

4. We present the first empirical model of mortgage default hazards using original data for a period covering the recent financial crisis.

5. We present the first empirical parameterisation of mortgage hazard rates that incorporates unobserved heterogeneity across borrowers and macroeconomic variables for the UK

1.3 Organisation of the Thesis

The remainder of the thesis is organised as follows. The next Chapter 2 outlines the existing valuation and prepayment/default literature. This review recognises the three prevalent research themes, identifies the significant features, characteristics and variables involved in their investigation and takes note of various mathematical and statistical methodologies used by researchers and industry practitioners alike.

Chapters 3 and 4 present the basic contingent claims modelling approach to developing the traditional general valuation model for mortgages and mortgage related products. Since the model has no closed form solution the overall procedure for its numerical solution is discussed.

In Chapter 5 the finite difference method used in order to reach a numerical solution for the model presented in Chapter 4 is considered.

In Chapter 6 the results of simulations involving a base set of key parameters are presented and discussed. Conclusions are drawn regarding the values of the different financial assets involved.

Chapter 7 presents a number of modifications to this traditional structural model. Their rationale, method of formulation, simulation and results are discussed.

Chapter 8 contains the empirical data analyses. These are based on three avenues of research found in the literature. Use is made of a proprietary dataset kindly provided by a financial institution upon which a number of reduced form models are constructed and comparisons made.

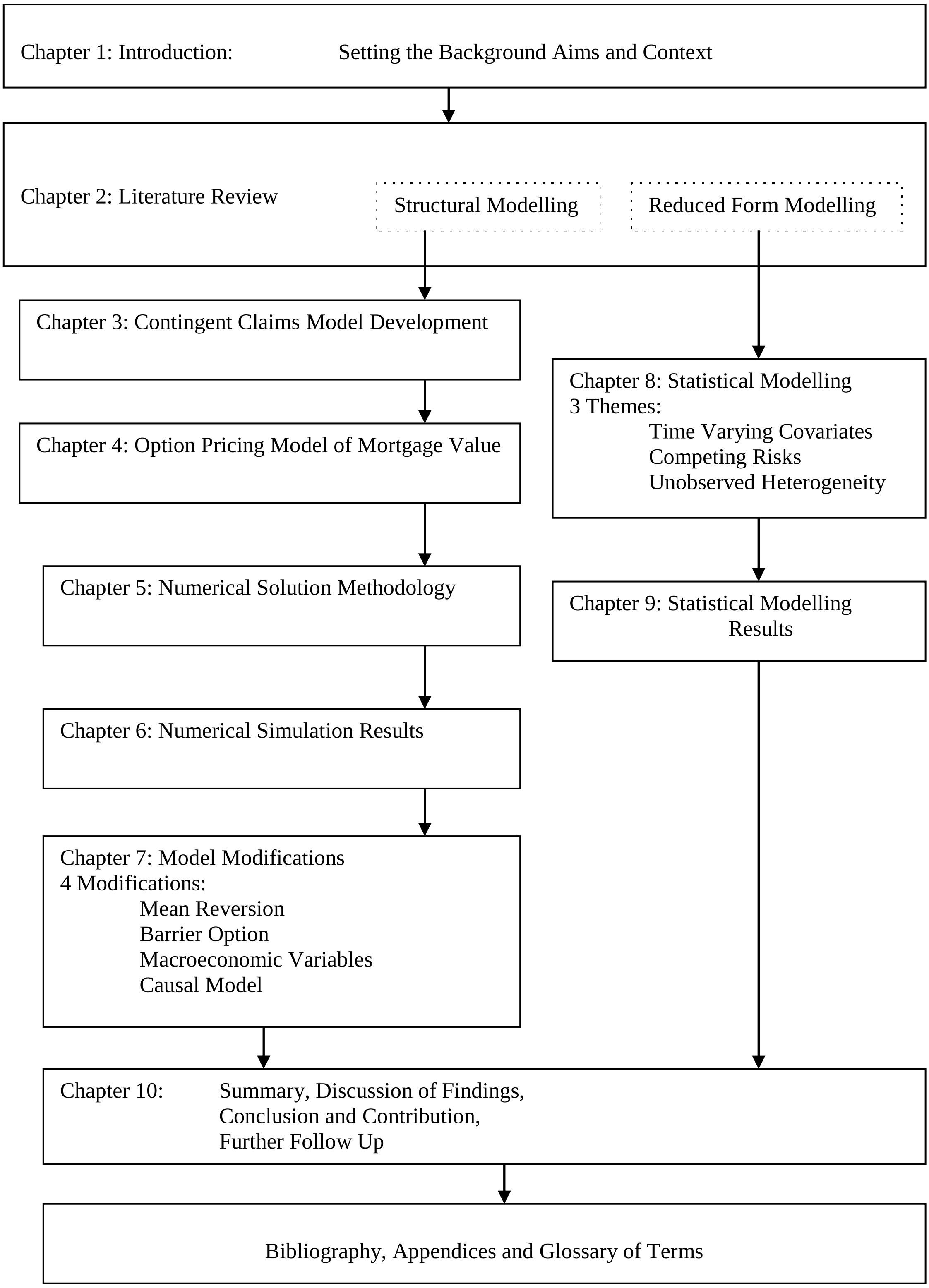
In Chapter 9 the results of the data analysis are illustrated with a discussion of the associated findings.

Chapter 10 concludes the thesis by identifying important contributions made coupled with suggesting areas for further follow-up research activity.

A number of Appendices and a Glossary of Terms used are also included. These are provided particularly in order to give simple numerical illustration to option related theoretical concepts as an aid to their understanding and present the associated mathematical derivations of model equations where their inclusion in the body of the text would detract from its overall readability.

A schematic of the structure of the thesis is also given to show visually how the chapters are linked.

Figure 1: Organisation of Thesis in Schematic Outline



Chapter 2

Literature Review

2.1 Introduction

There is a rich literature covering research into both mortgage default and prepayment. The present chapter provides an overview of the development of an ever growing body of knowledge that constitutes the contingent claims approach to mortgage valuation modelling and that encompasses these competing risks of termination. Through an initial chronological examination of the earliest literature – motivated by the reviews of Quercia and Stegman (1992) and Vandell (1993) – followed by a themed examination of the more recent literature – motivated by the reviews of Kau and Keenan (1995) and Leece (2004) – the costs of mortgage default and prepayment are examined and some of the theories proposed for describing default and prepayment behaviour are presented, in so doing the primary determinants and common research models are discussed. This chapter also identifies areas for future research and rationalises the choice of approach to the valuation of mortgages, the determination of mortgage default and prepayment rates and mortgage related products used in the following chapters of this thesis.

There now is an extensive and growing body of literature on Mortgage Performance Modelling[3] as a research activity. Consequently this offers a variety of

[3] Lenders who issue mortgages monitor their performance since they would like to know the value of the future cash flows that will be received as a result of borrowers making scheduled monthly payments or otherwise. The value of the mortgage is not simply the time value of these payments, since borrowers may terminate the mortgage prior to its maturity thus halting the projected cash flows. Lenders model the performance of mortgages they issue as their value also derives from the evolution of the global economy via underlying house price and term structure of interest rates.

different ways in which the mortgage performance modelling literature may be reviewed.

For instance, the research may be purely theoretical or empirical data driven or it may be classified as based upon either Residential Mortgages or Commercial Mortgages or indeed may be based upon pools of such mortgages forming Collateralised Mortgage Obligations (CMO) and other forms of Mortgage Backed Securities (MBS).

As will be exemplified in this literature review, the development of our current understanding of Mortgage Performance Modelling especially in the context of default risk has been influenced by studies of both Residential and Commercial Mortgage default experience. As discussed in Vandell (1993), in general historical terms, Commercial Mortgage default modelling has paralleled Residential Mortgage default modelling albeit lagging behind by about a decade or so.

2.2 A Time Line of Mortgage Performance Modelling

According to Quercia and Stegman (1992), the extensive body of research made until the beginning of the 1990's can be divided into a first-, a second- and a third *generation* of studies coupled with the *perspective* taken by the literature stream. More recent developments in the pre-US Mortgage crisis literature can be found in Leece (2004).

2.2.1 First Generation Studies

First generation models prevail from the 1960's until the early 1980's where modelling is viewed from the perspective of the individual mortgage lenders.

These very early studies had the purpose of assisting individual lenders in predicting the default probabilities of their borrowers.

Examples of such studies include those beginning with Jung (1962); Page (1964); Von Furstenberg (1969, 1970a, 1970b); and subsequently with Findlay and Capozza (1977) and Vandell (1978). Each of these studies contributes the very first insights into Residential Mortgage default risk. These studies attempt to explain the default outcome by providing evidence that demonstrates the importance of loan characteristics at origination, such as level of borrower equity and interest rate premia using aggregate data.

Jung (1962), despite a lack of formal economic theory, is the first to identify a positive correlation between the loan-to-value (LTV) ratio and default risk using aggregate data from U.S. savings and loans institutions.

Page (1964) and Von Furstenberg (1969) are among the first to use regression techniques to provide empirical evidence on the importance of LTV ratio at origination, mortgage term, amount of home equity at origination in default prediction.

Subsequently Von Furstenberg (1970a, 1970b) finds that home equity at origination is the most important predictor of default risk – default rates increase by up-to seven times if LTV ratios are increased from 90% to 97%. His findings also included evidence that default risk increases with the age of the mortgage up-to the loan's third or fourth year after which it declines.

Findlay and Capozza (1977) were the first to apply the financial option pricing theory of Black and Scholes (1973) and Merton (1973) in the mortgage arena as a basis for their estimation of the value of Mortgages. Suffering from limitations such as only focussing on the risk at loan origination, not giving full consideration to the time frame of the mortgage and only handling Fixed Rate Mortgages (FRM), these early studies using primarily data driven approaches typified by what may be termed *statistical* mortgage modelling techniques, can be seen as somewhat narrow.

However, they did provide the first indicators of the need to investigate the importance of borrower and property characteristics; the impact of borrower and property attributes receiving less consensus amongst the researchers of the time than mortgage characteristics. Some of the disagreement on the effect of borrower and property characteristics on default may be illustrated with reference to variables that reflect the fluctuation of household income, for example: Von Furstenberg (1969)'s results were indicative of household income having a negative and significant impact upon default risk. Attributing this effect to the correlation between income and LTV ratio Von Furstenberg concluded that household income by itself is not deterministic of default. Herzog and Early (1970), in testing a related variable, the debt-to-income ratio at origination also concluded that this had no significant effect on default. Although confirmed by Morton (1975) and Sandor and Sosin (1975) separately, a different conclusion was reached if the debt-to-income ratio was measured categorically (Williams, Beranek and Kenkel, 1974). Unlike Herzog and Early (1970), Williams et al. (1974), found that all else being equal, a borrower with a debt-to-income ratio higher than 30% is more likely to default than other borrowers.

In summary, this period of the early literature on mortgage termination is characterised by separate analyses of variables that influence default being undertaken in lieu of formal theory. Variables included in these analyses were taken from mortgage loan characteristics at origination, property characteristics, borrower characteristics and economic conditions. Researchers made no attempt at this time to provide a theoretical basis for borrower behaviour at the time of mortgage termination. However, they still made significant contributions to the line of research by identifying and testing candidate variables from which to determine mortgage termination. Thus, the evolution of mortgage default in the US as characterised by Quercia and Stegman (1992) typology begins here with the earliest work on default and delinquency risk from the 1960's onwards focussing from a lender's perspective upon simple correlations or empirical regression models capturing, at the origination of the loan, the characteristics of mortgages (LTV ratio, interest rate and mortgage term) and of the borrower (location, marital status, employment situation) that may be influencing later default. Compared to studies of default, at this time however there is little mortgage termination research that focuses primarily upon prepayment behaviour. Research into prepayment begins to flourish with the second generation of studies, when option theory is thoroughly established in the analysis of mortgage terminations.

2.2.2 Second Generation Studies

Approximately midway through the time period of first generation studies second generation models begin to appear; emerging in the mid 1970's they tend to prevail between 1980 and 1990 where modelling is viewed from the perspective of the individual borrower.

Here, the default decision of a borrower maximising utility and net worth was in focus. These studies are firmly rooted in the economic theory of consumer behaviour and constitute the basis of the current state of theory.

Option pricing theory applied to mortgage termination by default and prepayment derives from this second generation of studies by considering default and prepayment as embedded options in a mortgage contract and the dominance of net equity and interest rate as important loan characteristics are reiterated. The research of these second generation studies started with the valuation of a mortgage's embedded options to default and prepay, these options being treated separately and investigated individually. For example the works of Dunn and McConnell (1981a, 1981b) and Buser and Hendershott (1984, 1985) analyse the prepayment option of Adjustable Rate Mortgages (ARM) as well as Fixed Rate Mortgages. Green and Shoven (1986); Hendershott and Van Order (1987) and Schwartz and Torous (1989a, 1989b) analyse mortgage termination through prepayment using aggregated data. In respect of mortgage default, Titman and Torous (1989) analyse the special case of Commercial mortgages where prepayment is prohibited through yield maintenance features and prepayment penalties. Brennan and Schwartz (1985) carried out other research investigating the timing of default. The early work of Epperson, Kau, Keenan and Muller (1985) examined the valuation of Fixed Rate Mortgages and the pricing of related insurance against default on such mortgages.

Insurance and guarantees against default is also investigated in the work of Cunningham and Hendershott (1984).

The first published work that jointly considers default and prepayment options in a pricing model was that of Kau et al. (1987) in the context of Commercial Mortgages.

All of the above imply that adequate pricing of mortgages calls for a thorough understanding of the behaviour of mortgage borrowers and that this requires models of default and prepayment risk. The first connection between economic theory and mortgage borrower behaviours in academic discussion about the motives that drive mortgage termination can be traced back to the studies conducted during this period by Jackson and Kaserman (1980), Campbell and Dietrich (1983), Foster and Van Order (1984, 1985) and Vandell and Thibodeau (1985). This connection was very well established by the time of the studies of Zorn and Lea (1989) and Cunningham and Capone (1990). Each of these six studies utilises the economic theory of utility maximisation and tests previous research results based upon the premise that borrowers rationally decide whether to carry out the mortgage contract or not in the course of maximising their wealth over time. Studies by Giliberto and Houston (1989) and Kau et al. (1992a, b) have also incorporated analyses of nonfinancial termination motivations such as changes in family status and employment, these terminations also increase utility, but unlike financial terminations, the gains are not directly observable. Jackson and Kaserman (1980) presented a study in which the two main competing theories of mortgage default[4] were empirically tested.

[4] The first of these approaches asserts that borrowers observe a strictly optimising behaviour, maximising financial gain or minimising financial loss associated with the continuation or termination

By explicitly formulating and testing these two competing hypotheses the conclusion of the study supported the claim that the '*Equity*' effects were clearly dominant, so confirming the adequacy of an optimisation model of consumer choice in the analysis of default decisions. These results were later confirmed in a series of other similar studies as outlined below.

Campbell and Dietrich (1983), in their multi-period model of consumer choice extended the work of Jackson and Kaserman (1980) on the importance of net equity in the borrower's decision to default. They find that both the loan-to-value ratio at origination and a contemporaneous loan-to-value ratio were found to have significant and positive effects on the default decision. This coupled with their other result that a borrower's income variability over time also had a significant and positive effect provides the first evidence of the importance of time varying covariates on mortgage termination.

Foster and Van Order (1984, 1985) estimated a default function in an options framework that estimated the extent to which the proportion of loans with negative equity could explain aggregate default rates in a pool of mortgages.

Vandell and Thibodeau (1985) used a set of disaggregated loan histories from a U.S. savings and loan association to evaluate the influence of both equity and cash flow on default. They concluded that on the average an expected negative equity of 10% would give rise to approximately 5% likelihood of default.

---

of the scheduled periodic payments inherent in the mortgage contract. This view is termed the '*Equity*' theory of default. The alternative approach which is normally referred to as the '*Ability to Pay*' theory, holds that mortgage borrowers tend to avoid defaulting on their mortgage loans if their current income stream is adequate to fulfil the periodic payment without it being too great a burden on their expenditure.

Option Theory[5] was first formally applied to the field of mortgage default by Foster and Van Order in 1984 by significantly extending Campbell and Dietrich's work of 1983. A straightforward application of models by Black and Scholes (1973) and Merton (1973) for mortgage terminations provides us with a '*frictionless*' model in the sense that there are no transactions costs[6] associated with either default or prepayment.

The frictionless model assumes that borrowers can pay off the outstanding mortgage balance without penalty in regard to prepayment. It also assumes that borrowers can purchase a new property of equal value to the defaulted one and enter into a new loan arrangement after immediately exercising the default option again without penalty or reputation costs. Given its assumptions the model predicts that borrowers will immediately exercise their options whenever the default option or the prepayment option is '*in-the-money*' and hence the name '*ruthless*' termination behaviour is used to refer to it.

Despite frictionless models being a good indicator of the ability to evaluate mortgage contracts, empirical evidence does not completely rule out the existence of significant transactions costs.

In the work of Foster and Van Order (1984) the significance of the lagged equity terms indicates that the default option is not exercised immediately.

---

[5] The contingent claims option theoretic model views default as a put option, allowing the borrower to sell to the lender the house for the value of the mortgage at the beginning of each payment period. Likewise prepayment is viewed as a call option allowing the borrower to exchange a sum of money for the mortgage loan anytime during the lifetime of the loan.

[6] Most options theoretic approaches to mortgage valuation define transactions costs as the costs of exercising the underlying options. This includes selling and purchasing fees, taxes and more importantly limitations on future access to credit. It is obvious that if such costs were not negligible then they would tend to reduce default rates.

Given a borrower with negative equity they contend that an event such as a divorce or loss of employment may be needed to trigger default. However since there was no empirical support for this contention they attributed the imperfect exercise of the option to the importance of the financial costs of the transactions.

In a subsequent analysis, Foster and Van Order (1985) found that borrowers do not exercise the default option consistently, even under zero transactions costs or with negative equity. They did however acknowledge what they described as the simultaneity problem. Specifically, they stated that

> ‘ … An owner with negative equity with a mortgage evaluated at par might not default because he forfeits the option of defaulting later … ’

In fact, whenever an option to terminate a mortgage loan is exercised, the borrower automatically gives away the possibility of exercising both early termination options in the future. As a result borrowers tend not to default immediately after the value of the house falls below the present value of the future loan payments, but only after the house value reaches levels that are below the value of the mortgage including both options to terminate the loan. Vandell and Thibodeau (1985) also address the issue of transactions costs and crisis events which not only includes the financial charges incurred but also possible relocation costs and loss of reputation caused by defaulting.

Evidence of the non-ruthless exercise of the prepayment option is also present in the literature though here the terminology is to refer to this behaviour as '*sub-optimal*'[7]. Some borrowers do not prepay their loans even when the contract rate (coupon) surpasses the more favourable rate at which it would be possible for them to refinance the loan in the market. Similarly other borrowers prepay their loans in adverse conditions in which the refinancing rate is higher than the contract rate, (see Green and Shoven (1986)).

In summary, the connection between option theory and mortgage termination behaviours provides the researcher with a theoretical tool to analyse mortgage default and prepayment incentives and outcomes. Researchers in mortgage studies calculate the extent to which the default (put) option and prepayment (call) option are in the money, based on which they are able then to predict when default and prepayment occur. This generation of empirical models derived from theoretical models emphasises factors influencing the borrower's decisions on their alternative payment choices where the chosen outcome maximises utility over time given the individual borrower's circumstances. Such models emphasise the *financial* aspects of the decisions via negative equity and borrower characteristics are excluded. Thus these *frictionless* models predict identical behaviour for borrowers with similar mortgages.

Much of the empirical literature of this generation explores evidence that the default decision is delayed upon reaching a sufficiently negative equity position rationalising such evidence by transactions costs and future credit restrictions.

[7] The term sub-optimal is used because despite having some appeal to the borrower it does not minimise the market cost of the mortgage, as given by frictionless mortgage valuation models.

The many studies of this period in time conclude that evidence on the existence of non-ruthless behaviours, as well as the importance and impact of non-option related '*trigger*' events or crises affecting income, such as divorce, ill health or loss of employment, when intersecting with marginal levels of equity may initiate moves from the conditions of delinquency with negative equity into those of actual default. This introduces a role for ability to pay measures, in addition to equity on mortgage terminations and makes testing of both option and non-option variables a necessity.

Overall, the theoretical premises set forth in this second generation literature constitute the basis for the current state of the theory. The examination of the default and prepayment decision as options and the central role of net equity position constitute the dominant view. The second generation literature also provides some evidence regarding the nature and importance of transactions costs and other borrower-related factors. Although there are conceptual justifications for the relevance of crisis events on the default decision, there is little direct empirical correlation provided over this period in time. In the third generation of research studies illustrated in the next section, these and other issues are considered further.

2.2.3 Third Generation Studies

Third generation research models start to appear in the late 1980's and prevail from 1990 through to the present day. Third generation modelling mainly represents a technological improvement on that of the second generation. The use of econometric techniques in this mortgage loan literature has evolved from the use of qualitative dependent variable models to the application of proportional hazard and multinomial logit models incorporating competing risks to estimate default probabilities. Further developments have recognised the potential importance played by unobserved heterogeneity, particularly when modelling behaviour that appears sub-optimal from the viewpoint of the options theoretic approach (Deng et al, 2000, 2002).

Third generation modelling is viewed from an institutional standpoint. The focus here is not only on estimation of default probability and the prepayment rates of fractions of pools of mortgages, but also as a measure of mortgage risk that better reflects lenders concerns – expected mortgage loss – rather than default rates, as in most second generation studies concerned with modelling individual borrowers default and prepayment decisions. This arises from the need to price mortgage credit to anticipate losses from default. The risk of losses in large loan pools and the resulting capital requirements for lenders were examined by Quigley and Van Order (1991), who found that termination of the mortgage by default or prepayment before maturity adversely affects the profitability of lending institutions and produces lower than expected yields on Mortgage-Backed Securities for investors.

The research work of the late 1980’s including that of Green and Shoven (1986), Quigley (1987), Clauretie (1987), Clauretie and Herzog (1990) and Giliberto and Houston (1989) as with virtually all the previous research based upon options models, tends to focus on one not both termination outcomes by ruling out the possibility of default when valuing the right to prepay, and ruling out the possibility of prepayment when considering default.

Green and Shoven (1986) and Quigley (1987) were the first to use Proportional Hazards Models in the loan termination literature to analyse mortgage prepayment behaviour.

The foreclosure costs faced by lenders which are an important consideration in estimating the expected losses resulting from mortgage default were studied by Clauretie (1987) who found foreclosure to be more likely in those U.S. states where costs are lower. In U.S. states where legal costs were higher, lenders appeared more willing to work with slow or delinquent borrowers and so concluded that foreclosure rates '*reflect the choice of the least-cost method of limiting losses*' by lenders when dealing with slow or delinquent loans. In a further study Clauretie and Herzog (1990) also found that the loss rate on defaulted loans increased with loan-to-value ratio.

Giliberto and Houston (1989) examined the role of crisis events and default costs in their presentation of a theoretical model of default, though they did not test their model empirically. It was not until a series of papers by Titman and Torous (1989); Kau et al. (1992, 1995) and Kau and Keenan (1995) that recognition of the importance of the competing risk nature of the interdependent relationship[8] between default and prepayment and the joint estimation of the two termination outcomes became theoretically established. Quoting from Kau et al. (1992):

> ' … Since prepayment and default substitute for one another, contracts with only one of default or prepayment provisions leads the borrower to behave differently when both are present. This substitution effect means that one cannot accurately value either the individual provisions or their interaction without both options being present … '

[8] This is because a homeowner who exercises a default option today gives up not only the option to default tomorrow but also the option to prepay tomorrow, the two options are interdependent and need to be jointly estimated.

This is because the probability of default or prepayment is a function of the extent to which the other option is ‘in-the-money’.

In more recent studies of residential mortgages Deng, Quigley and Van Order (2000) and Downing, Stanton and Wallace (2005) have extended their earlier studies beyond simply prepayment to now include default. They have found that whilst default is a rare event among prime mortgages it is important through its correlation with prepayment. They echo the importance of joint prepayment and default by demonstrating the statistical significance of the joint options by finding that the variables that proxy for the prepayment option are significant for forecasting defaults and that the variables that proxy for the default option are significant in forecasting prepayment.

2.3 Methodological Issues in Third Generation Models

In the main however, third generation studies contribute fundamentally to an identification and exploration of certain methodological issues which will now be examined in a thematic way within the overall chronological structure presented herewith.

In particular we consider five themes. These are the use of different *measures of default risk*, for example expected mortgage loss that better reflects institutional concerns; the use of different *modelling techniques and solution methods* such as lattice approaches; finite difference methods and simulation; *estimation techniques* such as proportional hazards estimation of default probabilities and finally with *expedients and limitations*, particular shortcuts and other limitations prevalent in the literature.

Methodological Issue: Measures of default risk. From the perspective of lenders and investors (both of which may be institutions) default rates are not an adequate measure of default risk. To fully understand mortgage risk it is not sufficient to simply estimate the likelihood of default, it is equally important to estimate the '*severity*' of a mortgage – that is the percentage of the unpaid principal balance that is lost in the event of default – this is because the proportion of pounds loaned that become losses in default varies by loan[9], a fact that is not captured when default rates are considered. Severity is in one-to-one relation with the amount lost and thus its estimation fully captures the expected loss on a mortgage conditional upon default. Thus a measure of expected mortgage loss is a better indicator of mortgage risk, since expected losses provide a more accurate basis for estimating mortgage insurance premiums. Studies that have analysed the determinants of both mortgage loss and default rates are those of Evans, Maris and Weinstein (1985) and Clauretie and Herzog (1990).

[9] Severity is a more common way of modelling mortgage losses because many of the cost components, such as lost interest and commissions are related to the size of the loan. Such loss estimates are required to accurately price loans and derivatives and set economic capital and loss reserves, since under the Basel II framework regulated financial institutions are increasingly being allowed to set their own capital requirements using internal forecasting models.

Methodological Issue: Modelling Techniques and Solution Methods.

A Residential or Commercial Mortgage is specified by its contractual terms and conditions at its origination, and its behaviour is then determined by the interactions of these contractual features with the external economic environment through the actions and behaviours of the borrower. The most prominent contractual features are its payment structure – the most common residential mortgages, Fixed Rate Mortgages (FRM), have a fixed contract rate and a fixed monthly payment, so that more amortisation takes place near the maturity of the loan. In contrast Standard Variable or Adjustable Rate Mortgages (SVR) have calculated payments given the contract rate, but this rate is periodically adjusted to reflect the prevailing market interest rate. This variability can be partly offset by having lifetime or periodic caps (or floors) on the amount of change permitted by any given adjustment and typically a '*teaser*' rate is offered. This means that the first year's contract rate is artificially set below the prevailing rate in the market at mortgage outset – and the possibility of an early termination through either prepayment or default. Commercial Mortgages on the other hand whilst being typical of residential mortgages in terms of their payment structure are often also characterised by having initial lockout periods in which prepayment is not permitted and in having a final balloon payment, as the term used in calculating the payments is greater than the actual life of the loan. Both types of mortgages may have prepayment and default penalties explicitly provided for in their contracts even when prepayment is allowed.

Given that the payment structure of mortgages is notable for the length of time to their maturity, then to model the economy as having a constant or for that matter a deterministic term structure of interest rates would be inappropriate. Given the future is uncertain all research consequently considers the two main factors affecting mortgages operates within a stochastic economic environment. The single most relevant feature is a stochastic term structure of interest rates, in addition to account for the behaviour of the mortgage's underlying real estate asset, the house price process must also be stochastic.

The most popular choice for the term structure of interest rates has been to represent the spot rate as evolving according to a single state process the Cox, Ingersoll and Ross (1985b) (CIR) mean-reverting process. Brennan and Schwartz (1985) utilise a two state process involving a spot rate and a long rate. However Buser, Hendershott and Sanders (1990) compare the one-state and two-state interest processes in the context of mortgages and conclude that the one-state form is adequate. A two-state process has the advantage of modelling more closely the actual term structure but at the expense of the amount of computation required to obtain a solution.

Typically the assumption is that the house price process evolves as a geometric Brownian motion and is use extensively in the work of Kau et al. (1995). Empirical studies of house price dynamics suggest however that house prices are poorly approximated by a geometric Brownian motion process, Meece and Wallace (1997, 1998). In this research the effects of mean reversion upon mortgage pool valuation is examined.

The CIR term structure process is used with a Poisson parameter for the frequency of refinancing decisions and a beta distribution for transactions costs. The authors conclude that house prices may instead be more consistent with a mean reverting process.

Mortgages are complex contracts to value and so recourse must be given to numerical solution techniques, of which there are two categories. The first is the forward pricing method and the second is the backward pricing method.

The main technique of forward pricing is the use of Monte Carlo methods. The value of the mortgage asset is just the expected present value of its future stream of payments, discounted at the appropriate rate. We can therefore randomly select a path of interest rates and house prices, determine the mortgage value and then repeat the process until the average value closely approximates the mortgage's true value. Studies of Schwartz and Torous (1989, 1991) have adopted this procedure. The trouble being however that the approach is not appropriate when rational termination is allowed. To calculate the contribution of early termination to a mortgage's value at any moment, we must know whether termination is to occur. But this depends upon the borrower knowing the future values of the mortgage which has not yet been considered in the modelled process when going forward in time. While time moves forwards to value a mortgage with termination, we must begin at maturity, the end, and work backwards.

With the backward pricing method, at the maturity of the loan, when the mortgage value is known explicitly, given the economic environment it is possible to solve for the mortgage value at the previous instant in time for all possible interest rates and house prices, so if we continue with this backward reasoning the desired value of the mortgage at origination can be obtained for the actual interest rate and house price. This approach is in the manner of dynamic programming, by working backward through time, successively calculating the value of the mortgage at all the intervening times conditional upon the then current economic environment until we reach the end, the present time we obtain the value of mortgage at its inception.

The possibility of early termination complicates this procedure since a ‘free boundary’ is created which gives the critical combinations of interest rate and house prices at any time that induce early termination, however this critical termination boundary can be calculated by the backwards method as was not the case when going forwards. The information thus needed to determine whether the borrower would terminate is now available. Studies by Epperson et al. (1985) and Kau, Keenan, Muller and Epperson (1992, 1995) have adopted this method which provides the most popular means in the literature for implementing the option based approach via a backwards solving method for the solution of the asset valuing Partial Differential Equation (PDE) derived in continuous time.

There is however an alternative approach, which is to work in discrete time and space and by doing so effectively work through a tree or lattice structure during valuation calculations. For example the works of Hall (1985); Chen and Ling (1989); Giliberto and Ling (1989) and Leung and Sirmans (1990) provide application of the lattice model of Boyle (1988) along with very favourable comparison of the lattice models estimates with other models in the literature of the time. More recently, the works of Nelson and Ramaswamy (1990) in developing a generalised lattice approach; Hilliard et al. (1996, 1998) who extend the lattice to include more state variables and Ambrose and Buttimer (2000) further expand this frame work to incorporate a third variable to handle path dependency problem that exists in modelling the previous delinquency history of a mortgage.

Whilst there are distinctions between the discrete lattice approach and the continuous PDE approach they are in fact very similar.

The fundamental valuation PDE cannot be solved analytically and so must be discretised using a numerical procedure of which there are essentially two types, explicit methods and implicit methods.

Explicit Method. At any particular time the method generates the mortgage value by considering each grid point in the interest rate and house price plane individually using nearby grid points in the plane from a previous time step, going backwards. *Implicit Method*. This solves for the value of the mortgage by using the whole grid of points simultaneously, using the whole grid from the previous time step, going backwards.

The discrete lattice approach is considered to be an explicit method for solving the PDE with the important criterion that such a lattice procedure's solution converges in the limit to the solution of the PDE as the finite time intervals are taken to be shorter and shorter.

The terms of a mortgage contract are not arbitrary but are set so that the value of the contract to the lender equals the value of the amount lent to the borrower. This being the case then the mortgage contract is said to be in equilibrium at origination.

The mortgage value at origination if this was not the case would allow of contractual arbitrage opportunities which are clearly not desirable. Now, since insurance is a passive asset it is only through this balancing contract rate that insurance can exercise any influence on the mortgage. Balancing for Fixed Rate Mortgages is treated by Kau, Keenan, Muller and Epperson (1992, 1993, and 1995).

The development of early termination models especially in recent times acknowledges the path dependence of these and other fundamental economic variables through the use of Time Varying Covariates. The most common approach for incorporating time varying covariates is through the Cox (1972) and Cox and Oakes (1984) proportional hazard models however this is an open issue that has not been adequately resolved. An alternative to models based upon survival analysis is the use of a Random Utility Model. Even though discrete choice theory has been used in the study of early termination, Green and LaCour-Little (1999) and Clapp et al. (2001), an explicit link has not been made to the random utility maximisation that underlies this decision. Since a borrower undertaking an early termination decision is attempting to minimise the cost of their mortgage, or alternatively maximise their wealth, a Random Utility Model is a natural framework.

Random Utility Models were first introduced by Thurstone (1927) in the context of psychological stimuli. A formal link to utility maximisation was made by Marschak (1960). Based upon the McFadden (1973) consumer choice model, most studies assume that mortgage borrowers have multiple choices in their choice set at each payment period during the life of the mortgage. A borrower at each decision-making period chooses one of three choices that maximise their utility over time:

continue with scheduled payments, or stop payments and default, or prepay the mortgage. McFadden’s random utility model was employed by Vandell and Thibodeau (1985) and Cunningham and Capone (1990) to derive the conditional probability of each choice and by the use of multinomial logit approach estimated covariates that impact upon the termination outcome.

More Recently McFadden and Train (2000) have shown that any random utility model may be approximated by a mixed logit model.

Thus making them attractive tools for a variety of questions in which choice is to be made between discrete alternatives. Bhat (1999, 2000) explored random coefficients models with repeated choice-making by individuals under analysis proposing an ordered logit model with random coefficients. Since mortgage payment is a repeated event that is made at regular intervals, a borrowers decisions are likely to be influenced to a large extent by unobserved influences. Since a number of previously identified studies have found that variation in mortgage borrower behaviour cannot be explained fully by financial dynamics alone, the use of random coefficients in samples with repeated choices may be especially appropriate. This would allow the researcher to capture the heterogeneity that arises from the behaviour of mortgage borrowers which shares many of the fundamental aspects of areas in which random coefficients models have been successfully applied.

Methodological Issue: Estimation Techniques.

The joint estimation of the two options as previously explained calls for competing risks models. Since default and prepayment can be viewed as hazards - Han and Hausman (1990) proposed a hazard rate model that allowed for correlated competing risks - then survival analysis or duration models can also provide a convenient analytical tool to analyse early termination. Such duration models have received extensive attention following the more recent works of Deng, Quigley and Van Order (2000) who contribute to the literature by making competing risks duration models useful in the overlooked and so rarely analysed aspect of the mortgage market, namely the existence of unobserved heterogeneity amongst borrowers.

Indicated previously, unobserved heterogeneity can be modelled through transactions costs both monetary and intangible. Such intangible costs as time spent searching for a new home or access to refinancing opportunities are believed to account for much of this heterogeneity. Huang and Ondrich (2002) apply competing risk duration model and conclude that failure to control for unobserved heterogeneity leads to a severe downward bias in coefficient estimates.

Schwartz and Torous (1993) estimated the joint hazards of early termination through the use of Poisson Regression on aggregate data.

Methodological Issue: Modelling Expedients and Limitations

A number of shortcuts are taken in the theoretical literature on modelling mortgages, their legitimacy in application being dependent upon the studies objectives. These shortcuts have included: the consideration of Default Free mortgages, which allows the dropping of default and as a result the removal of the house price process, for example McConnell and Singh (1994) and modelling mortgage payments as requiring continuous rather than periodic payments, (see Chen et al. (1992); Pavlov (2001) and Liao et al. (2008)). Here the valuation calculation is not greatly affected although it does mean that default as well as prepayment is now considered to occur at any time.

A number of studies by the authors Schwartz and Torous (1989, 1991, 1992) dropped endogenous termination but included exogenous termination where the rule for termination included optimal termination.

Finally there is the assumption that early termination occurs in an effort by the borrower to minimise the costs of the mortgage even though there are reasons to terminate the mortgage when the strict observation of minimisation suggests that one should not – suboptimal termination – and likewise there are reasons not to terminate the mortgage when the strict observation of minimisation indicates that one should – transactions costs.

2.4 More recent developments

This segment of third generation modelling includes the more turbulent recent decade between the year 2002 and the present time, 2012. In this decade's studies, modelling is still viewed from an institutional standpoint. Such modelling continues to further amplify the intuitive empirical findings previously identified - namely, the respective roles of net equity, the evidence for borrower effects and transactions costs in default risk - but in addition focuses also on financial regulation and stability in the Mortgage Backed Securities (MBS) Markets[10] particularly in the aftermath of the US Subprime mortgage debt crisis of 2008. Studies on this particular aspect by Gerardi et al, (2008, 2009) and Bajari et al, (2008) make clear the neglected factors

of lending quality and credit constraints that do not find an obvious place in the utility maximisation framework underlying the pre-crisis second generational options theoretic models.

Other characteristics that are considered include the following:

- *The Mortgage Loan Characteristics* (see for example Azevedo-Pereira, Newton and Paxson (2002 and 2003) who apply option theory to specific mortgage products in the U.K.);
- *The Property Characteristics* Downing, Stanton and Wallace (2005) find evidence in support of differences in house price dynamics across different regions of the U.S. being an important source of heterogeneity between mortgage pool performance, thus confirming that it is necessary to include house price as a factor in MBS pricing in order to capture the effects of borrower default;
- *Borrower Related Factors* Bennett et al. (2001) found strong evidence that poor credit history as well as high current loan-to-value ratio significantly reduces the likelihood of refinancing. Kariya et al (2002) allowed borrowers to act differently within pools of mortgages. Kalotay et al. (2004) modelled

[10] A mortgage related financial derivative is a Mortgage Backed Security. This product derives its value from sets of mortgages of similar characteristics being grouped together and representative bonds backed on this pool - where the cash flows have been combined (securitised) - to form a more desirable debt instrument are sold to investors.

the full spectrum of borrower behaviour using the idea of refinancing efficiency. Longstaff (2004, 2005) using a multifactor term structure approach introduces the notion that a borrower's financial position affects the rate at which he/she can refinance bringing borrower credit into the analysis by adding a credit spread to the costs of prepayment.

- *Transactions Costs, Crisis Events and Expectations* Kelly and Slawson (2001) introduce prepayment penalties into the borrower decision making processes. Kau and Slawson (2002) demonstrate the flexibility and adaptability of an option theoretic model by allowing different frictions such as fixed and variable transactions costs, and sub-optimal and non sub-optimal termination.
- *The Double-Trigger Hypothesis*. This is a prominent concept in the mortgage default literature of this period. Note that a thorough early exposition of the double trigger model is expressed in the work of Elmer and Seelig (1998, 1999) with more recent application to the US subprime mortgage market by Bajari et al (2008). This theory views negative equity as a necessary condition for default but it also attributes default to the joint occurrence of negative equity and a borrower related triggering event such as unemployment or divorce that causes an adverse 'shock' to a household's finances. Gerardi et al (2008) and Foote et al (2008, 2009) specify reduced form models based upon frictionless and double trigger theories. In these studies the frictionless theory was found to be excessively sensitive to changes in aggregate house prices and predicts a far too strong rise in default

rate. By contrast the double trigger theory was found to be consistent with the evidence. The authors' economic reasoning suggested that default rates have increased roughly in proportion to the number of borrowers who experienced ANY level of negative equity as predicted by the double trigger theory, whereas the frictionless theory predictions are based upon the number of borrowers experiencing EXTREME levels of negative equity this having increased by much more post crisis than actual default rates.

- *Mortgage Backed Securities Markets*. Increased risk in this market, as represented by increases in residential mortgage delinquency and foreclosure rates - indeed when interest rates drop, a wave of repayments follows, c hanging the financial characteristics of the MBS as mortgage proceeds are passed directly to bondholders, shortening the duration of their portfolios - has had a significant effect on the financial markets and the economy as we have recently witnessed.

The majority of third generational empirical studies focus on the prime lending market in the US rather than subprime or non-conforming loans. However, there is a growing body of literature focussing primarily upon subprime mortgage loan performance, with analysis again being to a large extent US specific (Alexander et al, 2002; Pennington Cross, 2003; Cowan and Cowan, 2004; Deng and Gabriel, 2006; Pennington Cross, 2006; Pennington Cross and Ho, 2006, 2010; Danis and Pennington Cross, 2008).

Econometric specifications and findings in the research of the subprime market tend to reflect the research into the behaviour of prime mortgage loans and finds that relative to the finding on prime loans, covariates have larger marginal effects on the default and/or repayment probabilities (Chomsisengphet and Pennington Cross, 2006). This literature emphasises the effect on default rates of originating loans from third parties (Alexander et al, 2002; Pennington Cross, 2002); how default rates vary by loan classification (Cowan and Cowan, 2004) and the influence of contract features such as prepayment penalties and reduced documentation (low or zero documentation in the US and its equivalent the self certified mortgage in the UK) upon the likelihood of default (Quercia, Stegman and Davis, 2005; Rose, 2008; Pennington-Cross and Ho, 2010). This relaxed approach to mortgage underwriting attenuates or overrides prudential lending criteria and introduces information asymmetry, with the lender knowing less and less about the borrower’s ability to pay and likelihood of default. As a consequence such reduced documentation contracts will have a positive effect on the likelihood of defaulting.

Kau, Keenan, Lyubimov and Slawson (2011, 2012) develop a reduced form model to assess subprime mortgage default risk over the period 1997-2008. They demonstrate the need to stratify by period of origination and that borrower behaviour changed in a statistically significant way, consistent with the conventional wisdom that underwriting practices underwent a substantive deterioration. These results raise pertinent questions about the extent of information asymmetries between borrowers and lenders and the efficiency of the mortgage lending process.

Kau, Keenan and Li (2011), explore a shared-frailty model providing a method for modelling survival data when the survival times were not independent. Their empirical results suggest that it is important to control for group level frailty (a stratification by period of origination) to account for the within-group correlation amongst individual mortgages. Differences in environment and in macroeconomic setting were also found to have an important influence on mortgage termination risks.

2.5 Credit Risk: Measures and Modelling under Basel II

Significant attention has been devoted to the subject of credit risk measurement by the international regulatory and banking communities fostering extensive literature in credit risk by both academics in finance and practitioners in financial market industry. This is characterised by the co-existence of two parallel credit risk approaches based upon a simple dichotomous rule of data availability leading to two streams of credit risk modelling that have key distinctions. There are (a) the prices observed from a credit market, (from which the *risk-neutral* default probability implied from the credit market data is found) used in the academic literature of corporate credit risk which has been inclined to study implied defaults, and (b) the direct measurements of credit performance (the direct observations of defaults, also known as the *Actual* or physical default probability in finance) which includes the popular industry practices of

- credit rating in corporate finance by for example Fitch, Moody’s or Standard and Poor’s and
- credit scoring in consumer lending by for example Experian or Equifax.

Both credit ratings and scores represent the creditworthiness of individual corporations, mortgage backed derivative securities and individual credit consumers (mortgage borrowers). Credit risk thus affects virtually every financial contract; therefore the measurement, pricing and management of credit risk have received much attention following the Basel Committee on Banking Supervision (BIS, 2001) reforms to the capital adequacy framework by introducing risk-sensitive capital requirements. In attempts to reduce the chance of systematic failure of financial institutions, the various capital requirements regulations of the Basel II accord have been adopted.

The New Basel Accord allows a bank to calculate credit risk capital requirements according to either of two approaches: a standardized approach which uses agency ratings for risk-weighting assets and an Internal Ratings Based (IRB) approach which allows a bank to use internal estimates of components of credit risk to calculate regulatory capital. This allows financial institutions (banks) to estimate the minimum amount of *'regulatory' capital* they are expected to hold subject to approval. The amount of capital that a bank must hold to cover for default differs for the types of loan held. For Mortgage loans, the regulatory capital cover for credit risk is given by the following expression:

To calculate the minimal capital requirements, *K*, under the Internal Ratings Based Approach:

$$\mathrm{K} = \mathrm{LGD} \times \left[\mathrm{N}\left[\frac{1}{\sqrt{1\text{-}\mathrm{R}}} \times G(PD) + \sqrt{\frac{R}{1-R}} \times G(0.999) \right] - PD \right] \times \left(\frac{1 + (M - 2.5) \times b(PD)}{1 - 1.5 \times b(PD)} \right)$$

Where,

R = Asset Correlation; the correlation between an individual loan and the global state of the world economy,

N[…] = the cumulative distribution for a standard normal variable,

G[…] = the inverse cumulative distribution for a standard normal variable,

LGD = Loss Given Default,

PD = Probability of Default,

M = Maturity of the loan,

b(PD) = Smoothed regression maturity function. The slope of the adjustment function with respect to M decreases as the PD increases.

Regulatory Capital = K x EAD

The specified value of R for mortgage loans is R=0.15

Financial institutions then aim to develop methods to estimate / build statistical models to predict the following components of their loan portfolio:

- PD (probability of default in the next 12 months);
- LGD (loss given default);
- EAD (expected exposure at default).

if they use the Basel II advanced internal ratings based (IRB) approach.

Modelling the probability of default PD, (the objective of credit scoring systems for some fifty years) has long been one of the objectives of credit research, and this review has been the subject of the thesis so far where the general consensus from the literature on the relative empirical performance of ‘structural’ and ‘reduced form’ default probability modelling, the relative accessibility of data and modelling software all combined to lead to a rapid expansion in the credit modelling literature in the mid-2000s.

Compared to either the industry practices as mentioned above or the academic approach of the implied default probability, the academic literature based on the actual defaults is rather small which may be largely due to the limited access for an academic researcher to the proprietary internal data of historical defaults.

However, modelling LGD, in the consumer debt market, is not something that has really been properly addressed. That is until the advent of the Basel II Capital Accord which allows banks the opportunity to estimate LGD using their own models with the advanced Internal Ratings Based (IRB) approach. What LGD modelling had been done was mainly in the corporate lending market. Since 2006, there have been some papers modelling consumer LGD, however this research is still in its infancy.

Qi & Yang (2009) use linear regression to model LGD in mortgages observing that LGD could be explained by the loan characteristics; the nature of the underlying property, and variables measuring the default, foreclosure and settlement process. The most important factor they found is the current loan-to-value ratio.

Leow et al (2012) model mortgage LGD by using the probability of repossession multiplied by a haircut model to predict LGD. They found that the two-stage model was more affective at accurately reflecting the LGD distribution.

The Basel regulations require lenders to use economic conditions as part of the model. After the recession of the late 2000s, regulatory agencies have focused a great deal of energy on incorporating macroeconomic sensitivity into credit models.

Reduced form models are naturally suited to this approach wherein the idea of employing survival analysis for building credit-scoring models was first introduced by (Narain, 1992) and then developed further by (Thomas *et al*, 1999 and Stepanova and Thomas 2002).

Thomas et al. (1999) compared performance of Exponential, Weibull and Cox's nonparametric models with logistic regression and found that survival-analysis methods are competitive with, and sometimes superior to, the traditional logistic-regression approach. They describe survival analysis as a means to build dynamic models since this readily allows the inclusion of macroeconomic variables as time-varying covariates (TVCs). Bellotti and Crook (2009) using the Cox proportional hazard survival model to model time to default for a large database of credit cards show that accounting for macro-economic indicators such as interest rate and unemployment index significantly affects default probability and find a modest improvement in predictive performance in comparison to a static logistic regression. However, a difficulty for including time varying macroeconomic variables for most empirical models is that until the credit crunch the UK had been enjoying a relatively uneventful economic situation for the last fifteen years, and that most data used in modelling was collected during that period.

To gain variations in macroeconomic variable values some data sets for modelling come from the recession during the 1990's and the relatively uneventful period after.. The improvements to LGD models by including economic variables are demonstrated by Leow et al (2012) who, using macroeconomic variables to predict

LGD found that while they were significant they had very little effect on improving the predictive performance of their model.

Whilst an extensive research study on the LGD is outside the scope of this thesis, the structural valuation models supported by the reduced form statistical modelling discussed contribute to the general body of credit risk research assisting financial institutions in their aim of using the Basel II advanced internal ratings based (IRB) approach.

The theoretical and empirical parts of this thesis therefore are linked: The individual mortgage value underlying the valuation of pools of mortgage loans needs to be accurately priced (illustrated by the structural modelling herein) and the likelihood of default needs to be accurately determined (illustrated by the reduced form modelling herein) if any losses on default are to be reliably quantified and bank regulatory capital put aside to mitigate these losses.

Running through the entire body of this reviewed literature, there are two distinct modelling strategies discernable, with the natural distinction between them being dependent upon whether an option-theoretic rational valuation, or *structural*, model is used or whether an econometric valuation, or *reduced-form*, model is employed. In short, structural models as presented in the literature have as primary focus pricing; they model option exercise events by linking them to the fundamental underlying dynamics of interest rates and the asset that is the collateral for the mortgage. In these models mortgage termination is the result of a strictly optimising behaviour of the borrower allowing the structural methodology to produce informative forecasts characterised by this economic environment.

However despite their flexibility to include many subtleties of mortgage contracts, it is true that these models find it difficult to give termination predictions that match observed behaviour.

In contrast, reduced form models are those that do not explicitly model the capital structure and assets of the borrower, but rather model borrower decisions based upon a set of exogenously specified explanatory variables. Via the functional form of the model, they provide a degree of flexibility to closely match historical data, however these models emphasise empirical estimation of the random timing of option exercise events and unfortunately, do not determine how the estimated parameters would fare in response to a change in the economic environment which leads to notable poor predictive performance out-of-sample as discussed by Downing et al. (2005). Regardless of this shortcoming however a reduced form model can provide an understanding of empirical performance of existing contracts and insights into their pricing dynamics.

By contrast with an extensive US on literature on mortgage default and a fairly modest contribution on mortgage delinquency, the set of UK studies on default and arrears is strikingly limited. The relative scarcity and limited quality of loan level micro-data sets in the UK has constrained the types and extent of analyses that can be carried out. Only a small subset of UK studies using various datasets adopts a disaggregated approach to explore the determinants of UK regional default and repayment difficulties.

For instance: Muellbauer and Cameron, (1997) use the CML (Council of Mortgage Lenders) Database; Cooper and Meen, (2001) use regional possession court orders and Gathergood (2009) uses BHPS (British Household Panel Study) data. Lambrecht et al, (1997, 2003) employ a proprietary dataset supplied by a mortgage insurance company. The earlier 1997 study extended the traditional option theoretic approach examining both equity and ability to pay variables influencing UK default finding the latter to have more of an influence on default than the former.

The later 2003 study examined similar influences on the timing of default using a hazard model. The study of Aron and Muellbauer, (2010) incorporates a typology of UK studies on mortgage arrears and default as undertaken throughout this period.

In summary, it is well established that the option-based contingent claims models provide connections between formal economic and financial theory and mortgage borrower behaviour. Although challenged, option related variables (specifically variables measuring the market value of the mortgage) remain to be explanatory in determining mortgage terminations. Along with variables reflecting transactions costs and crisis triggering events, option related variables continue to draw attention from researchers and are included in estimation of mortgage terminations.

Competing risks models are called for when estimating mortgage terminations because it is well known that default and prepayment are interdependent and need to be jointly estimated.

The existence of borrower heterogeneity should be accounted for if one needs to correctly draw inference on the effect of explanatory variables on either option.

Both structural and reduced-form models have the same goal that is to account realistically for all the embedded options in mortgage contracts and as Kalotay et al. (2004) comment there is evidence for the use of both types of model in practice. As mentioned, structural models have the flexibility to include specific mortgage features, frictions and borrower heterogeneity and reduced-form models have the flexibility to match historical data but given the functional form used may not perform well out-of-sample.

2.6 Concluding Remarks

This chapter has presented the key theoretical and empirical approaches adopted in the mortgage loan performance literature. and positions the research of this thesis in relation to this work. The discussion facilitates the identification of the fundamental influences upon both default and prepayment behaviours which informs both the theoretical modelling and simulations of the next chapter and the empirical model that follows in a later chapter.

This body of work establishes the requirement for theoretical options based valuation models that utilise a contingent claims and competing risks framework and empirically for estimating a statistical model of default and prepayment behaviour with competing risk and unobserved heterogeneity that incorporates both an option theoretic specification and includes variables that impact upon affordability or reflect exogenous ‘shocks’ using a loan level dataset.

The next chapter, Chapter 3 opens with the theoretical modelling to be undertaken and this will be of the option-theoretic structural form. The choice of the processes for the underlying two state variables that model the economic environment, consistent with recent literature will be presented and derivation of the asset valuation PDE using standard hedged portfolio arguments will be given, provision is made for its standard solution and discussion is given to a number of adaptations.

In a subsequent chapter a reduced-form model is derived from loan level payment history data. Modelling being consistent with the three themes prevalent in the literature, dealing with the time varying nature of explanatory covariates, the competing risks nature of default and prepayment and the unobserved heterogeneity of the borrower.

Chapter 3

Mortgage Valuation using a Contingent Claims Analysis: Overview and Assumptions

3.1 Introduction.

In this Chapter the foundations of an option-based approach for the valuation model used in this thesis is presented and applied to value a Fixed-Rate Mortgage. Specific consideration is given in the structural model so derived to the embedded options of default and prepayment which are determined endogenously. Using numerical procedures, a number of risk factors and how they impact on the value of the FRM are discussed. These risk factors include for example, the volatility of interest rates and house prices and the transactions costs of the borrower in exercising the prepayment and default options; default and prepayment penalties. Such costs are considered to be a constant portion of the unpaid balance on prepayment or a constant portion of the value of the underlying property on default termination (Stanton (1995) and Kalotay, Yang and Fabozzi (2004)).

The model leads to the solution of a partial differential equation (PDE) with two state variables, the interest rate and house price. A closed form solution to the PDE is not possible, so numerical methods are generally used. Due to the problem of the free boundary posed by the early-exercise property of the American-style Call option to prepay at any time and the Compound European-style Put option of default only on payment dates an explicit finite difference method is used (Azevedo-Pereira, Newton and Paxson (2000, 2002, 2003)).

Whilst this method suffers to some degree from instability in the solution due to restrictions placed on the time steps. Following this solution, modifications are incorporated into this PDE, in an attempt to address gaps in the literature. A discussion of results is given after numerical illustration with benchmark parameters similar to Kau, Keenan and Kim, (1994).

3.2 A Contingent Claims Model of Mortgage Termination.

The option based approach is derived from the contingent claims analysis of Cox, Ingersoll and Ross (1985a), which models derivative securities based upon a PDE. The use of Contingent Claims Analysis in evaluating mortgages uses two assumptions:

Firstly is the assumption of Perfect Capital Markets, that is, a market devoid of any arbitrage opportunities[11] (Arbitrage reasoning is based upon the ability to costlessly replicate the derivative asset under consideration) and secondly, the principle that a borrower acts in a rational impersonal way to minimise the market cost of their mortgage. This implies that borrowers with identical houses and identical mortgages would always terminate their mortgages in identical circumstances.

In order to calculate the probability of default termination another key point needs to be borne in mind which is, there must be prior pricing of the mortgage components. It is only by a comparison of these components that a borrower can decide on whether to continue to pay, prepay or default on their mortgage.

[11] Arbitrage is the process of attempting to profit by exploiting price differences of identical or similar financial instruments, on different markets or in different forms.

To motivate the remainder of this section in order to establish the value of the components of a mortgage contract, a more specific look at mortgages is undertaken. To develop the model the main focus is at the level of the mortgage contract. A mortgage is specified in terms of the features of its contract and its behaviour is determined by how these features interact with various aspects of the economic environment. The two prominent contractual features are its payment structure and the possibility of early termination by prepayment or default.

There are two basic types of mortgage contracts: Fixed Rate Mortgage (FRM) Contracts and Adjustable or Variable Rate Mortgage (ARM) Contracts. These categories of mortgage contract can be distinguished by their individual payment schedules. These differ in the calculation of the contract rate[12] and the degree of Amortisation.[13] The payment structure of a mortgage is unique in terms of the length of its term to maturity. This is very long (up to 25 years in UK and 30 years in US).

3.2.1 Modelling the Term Structure of Interest Rates

The valuation of any financial asset is an attempt to adequately relate its future cash flows to the corresponding discounting factors. The economic environment exerts a decisive influence over the evolution of these elements. Faced with the uncertain evolution of both future payments and the discount factors, we face a source of interest rate risk. As a result, the evolution of the interest rate variable that characterises these sources of risk needs to be considered.

---

[12] The Mortgage contract rate: For a Fixed Rate Mortgage this is the fixed annual rate of interest charged on the mortgage and is used in determining the fixed monthly payment. For an Adjustable Rate Mortgage this rate is adjusted periodically to reflect the prevailing interest rate.

[13] Amortisation refers to the gradual elimination of a liability in regular payments over a specified period of time. Payments include both Principal and Interest.

Whenever the environment is assumed to be uncertain, the valuation function needs to take account, in a probabilistic way, the possible states of nature that can take place at any moment in the future. The classic Net Present Value (NPV) calculation is inadequate since neither the cash flows nor the rates that will be necessary to discount these cash flows are known with any certainty. Likewise since there is a long term to the maturity of mortgage contract it is inappropriate to use a flat term structure of interest rates. This means that it is inappropriate to model the economy as having a constant or even deterministic term structure, it necessary to make use of a stochastic term structure.

Interest rates are natural state variables in any mortgage valuation model. In the UK it is the norm that a mortgage is expected to be outstanding for up-to 25 years. Therefore potentially all interest rates up to this maturity can assume the role of state variables. However, for tractable solutions it is assumed that all interest rates are driven by one, possibly two, exogenous rates.

Whilst different interest rate processes have been used the most common, and dominant, specification is that proposed by Cox, Ingersoll and Ross (CIR, 1985a, b). It is a mean reverting process with local variance proportional to the level of the spot rate, which always prevents interest rates from reaching negative values. This dominance may be explained by the research done by Buser Hendershott and Sanders (1990) who conclude that the bulk of the yield curve variation is accounted for by a single factor. The CIR models' overall performance in mortgage valuation frameworks has been analysed (Archer and Ling 1995) with good comparative results.

In this thesis it is assumed that the term structure of interest rates is fully captured by this mean reverting process. Besides the term structure of interest rates, the other main factor that is necessary to consider in characterising the economic environment is the stochastic process followed by the real asset that underlies the mortgage contract.

3.2.2 House Price Dynamics

Option theory based mortgage valuation models assume that house prices follow a standard log-normal (geometric Brownian motion) process to represent the evolution of the property value over time (Cunningham and Hendershott, 1984; Epperson et al, 1985; Kau, Keenan, Muller and Epperson 1990, 1992, 1993, 1995). The value of the asset is assumed to evolve at a constant rate that is continually disturbed by a stochastic factor.

In this thesis it is assumed that the evolution of house price follows a standard log-normal process in the first instance. However, there has been a body of empirical evidence (Meese and Wallace (1997, 1999); Englund, Gordon and Quigley (1999)) that seems to suggest that the evolution of house prices may be more consistent with a mean reverting process. Microeconomic theory also provides support for the notion that the price of a good (in this case the property) in a competitive market is mean reverting, due to supply and demand responses to prices. The issue of the evolution of house price following a mean reverting process will be addressed subsequently.

3.2.3 Mortgage Holder Behaviour

It is commonly held that Mortgages can be viewed as ordinary debt instruments with various options attaching to them. There is a default option that is viewed as a put option and a prepayment option that is viewed as a call option. The purchaser of a property via an ordinary mortgage thus receives a "financial package" that consists of a loan with two embedded options:

- An American Call Option to prepay the loan at any time[14], or in other terms, an option to acquire the full set of rights over the house by paying-off the remaining debt; and
- A European Compound[15] Put Option (a succession of European Put Options) to default on the loan at any payment date[16], or in other terms, the house is sold back to the lender in exchange for eliminating the mortgage debt.

When modelling mortgage termination risk, it becomes necessary to model competing risks: the continual choice for the borrower to maintain payments, (that is, to pay the next loan instalment, by doing so the borrower is then continuing to honour their contract with the lender and will then be faced in the next month with the choice again as to whether to continue, prepay or default on the loan), prepay the loan or default on the loan.

---

[14] The accumulated interest due upon prepayment is rising at the contract rate during the whole of the time between payment dates and so the borrower can exercise their right to prepay at any time.

[15] They are compound in the sense of Geske (1979) in that at each payment date prior to the last one, the borrower either defaults or by making the scheduled payment purchases a new option to default at the next payment date.

[16] It only makes sense for the borrower to exercise their right to default at a payment date since earlier exercise confers financial benefit (in the time value of money) to the lender and the borrower may freely enjoy the benefits of homeownership until such time.

Although these choices for the borrower are distinct, they are not independent, since choosing to default at a particular time, means the loss of the option to prepay (or default) at a later date.

This makes the options to default and prepay serve as substitutes (Kau, Keenan, Muller and Epperson, 1992) and so any change in the value of one affects the value of the other. Thus the borrower holds a joint option to terminate their mortgage. To better understand the economic behaviour of home owners it is therefore crucial to appropriately model prepayment and default risks.

In modelling prepayment risk, option pricing theory indicates that a borrower will terminate their mortgage contract by prepayment when the value of the mortgage exceeds the outstanding balance. When mortgage interest rates fall relative to their contract rate borrowers have an incentive to prepay in order to refinance, provided of course, the LTV ratio still meets lending guidelines. The term structure of interest rates is thus the primary determinant of mortgage prepayment.

However, some borrowers do not prepay their loans when the contract rate surpasses the rate at which it would be possible for them to refinance the loan in the market. This type of prepayment is termed *suboptimal* prepayment, because it is in part enforced through the borrowers own personal circumstances (these micro-level personal circumstances involve the likes of job change involving relocation, or divorce or redundancy involving reduction in income or even a windfall increase in wealth)

In modelling default risk, option pricing theory indicates that a borrower will terminate their mortgage contract by defaulting when the value of the mortgage exceeds the value of the underlying property. But this view of borrower behaviour ignores the observation that even if the default option is deep 'in-the-money', the borrower may not exercise it in order to avoid a bad credit rating and so preserve future access to credit.

The borrower's current equity position in the mortgaged property as given by the loan to value LTV ratio and the size of the borrower's mortgage payment obligation, relative to his disposable income as given by the payment to income PTI ratio are then the two primary determinants of Mortgage Default. These ratios encapsulate the equity effects and the ability-to-pay effects identified in the literature review of Chapter 2. Prepayment and default can occur for reasons unrelated to interest rate and house price movements.

3.2.4 Departures from Perfect Capital Markets

The assertion that borrowers observe a strictly optimising behaviour, maximising their financial gain, or minimising their financial loss resulting from a ruthless application of their rights to the termination of the mortgage contract is not strictly observed in practice.

This type of termination is termed suboptimal, because it is in part enforced through the borrowers own personal circumstances. These non-financial terminations are driven by exogenous crises, '*shocks*' or '*trigger*' events which may involve the likes of job change involving relocation, or a change in family status involving divorce or a reduction in income through unemployment or redundancy. As these events are often involuntary, unpredictable and not as a result of wealth maximising criteria then such terminations frequently occur when the options are 'out-of-the money'.

Recall from the literature review of Chapter 2, the Double Trigger Hypothesis, where the relative importance of two key drivers of mortgage defaults: negative equity and borrower illiquidity were introduced. Under this hypothesis default is attributed to the joint occurrence of negative equity and a borrower related triggering event causing an adverse '*shock*' to household finances.

The '*option theoretic model*' of mortgage default is traditionally interpreted as implying that borrowers should default if and only if they have negative equity in their home. However, as the numerous studies previously mentioned show many borrowers with negative equity do not default; and, conversely, default is often associated with 'shocks' such as divorce, ill health and unemployment that trigger an inability, by the borrower in crisis, to maintain income levels sufficient to continue to pay their mortgage obligations as well as those shocks affecting the level of equity in the home from changes in interest rate and house prices.

Since the option theoretic model deals endogenously with shocks relating to equity level, attempting to explain observed practice means attention is given to shock events triggered exogenously. The shock of divorce leads to a sharing of resources between two households resulting in a reduction in income per borrower. Likewise periods of ill health particularly if protracted, lead to a reduction in income and unemployment leading to the loss of borrower income.

Work by Ford et al. (1995, 2004) and Munro (2000) is beginning to develop an understanding of the strategies which people adopt to cope with such exogenous shocks. Moreover, properly understood, the option theoretic model does not imply that negative equity alone is sufficient for default. By defaulting today, a borrower gives up the option to default in the future; as a result, even with negative equity, a borrower might prefer to wait and see if house prices recover (Kau et al., 1994).

The standard way of reconciling the option theoretic model and what is observed in practice (through data collection and subsequent analysis) is to introduce *transaction costs of defaulting*, such as moving costs, reputation costs (e.g., lost access to credit), and stigma. But such costs can be difficult to identify empirically, often being represented as a percentage of the unpaid balance on prepayment and as a proportion of the property value on default.

The possibility that an individual will not default when a strict examination of house price and mortgage terms indicates that they should, can thus possibly be attributed to transaction costs. These costs include the tangible financial costs of option exercise (prepayment and default penalties) as well as those that stem from the costs incurred as a result of non-financial termination such as access to further lines of credit or loss of reputation. As identified in the literature review of second generation models the role and importance of transaction costs is subject to some debate.[17]

---

[17] In that borrowers do not behave ‘ruthlessly’, deeply in-the-money default is exercised even if transaction costs are negligible and it is not possible to reject the hypothesis that transaction costs matter to findings of differences in loan loss severity.

In addition whilst the individual costs of termination are not generally observed they are typically used to explain the residual between the idealised model results and actual data. Buist and Yang (1998) observe that the gap between theory and practice can be attributed to a number of causes:

'… … (a) The prepayment and default rate predictions of the theoretical models exceed observed rates … …' which may in part be explained by the effects of suboptimal termination;

'… … (b) The theoretical models have mostly relied on just two stochastic processes, interest rates and house prices … …' ignoring income or other more complex term structures or different house price dynamics;

'… … (c) Mortgage underwriting constraints, key variables like consumer credit scores and other factors … …' do not fit easily within the theoretical framework.

The increase in mortgage default rates in recent years has created renewed interest in these factors driving mortgage default because there has been an increase in the number of subprime loans originated after 2003/4, due to lax mortgage underwriting standards. This has increased the number of borrowers who are more susceptible to unemployment and negative home price shocks. These borrowers have little savings they could use to cushion themselves against such unemployment and negative home price shocks.

More recent studies have drawn conflicting conclusions as to which of these two factors have accounted for most of the variation in the resulting mortgage default rates.

As Elul et al (2010) conclude there are important policy implications for how best to help home owners in difficulties depending on which factor dominates - if negative equity dominates, then US government programs that reduce the overall principal might be beneficial. By contrast, if unemployment shocks should dominate, reduction in payments (or subsidisation of mortgage payments) might be the better US government policy. The second conclusion is that during the last two decades default experience has been significantly influenced by changes in regional rates of unemployment and as a result there is a continuing need for geographic diversification in mortgage default risk.

The empirical data analysis of this thesis (Chapter 8), considers the nature of transactions costs in the options theoretic model in the form fairly typical industry figures applied on a sliding scale over a short period of time along with certain other borrower characteristics, loan characteristics and shocks/trigger events such as economic conditions in addition to the financial option variables.

In this thesis an option-theoretic, structural model is developed on the basis of a number of fundamental assumptions. It assumes borrowers behave '*ruthlessly*' in defaulting to maximize their wealth (minimise costs) and this coupled with the further modelling assumption of being '*frictionless*' - in the sense that the borrower can exercise default without penalty - means that the model predicts an '*identical*' default behaviour for all borrowers with similar mortgages.

However as mentioned in the literature review of Chapter 2, there is in practice evidence to demonstrate '*non-ruthless*' borrower behaviours.

This delayed default behaviour has been explained in part by transactions costs, future credit constraints and the impact of '*trigger events*' or crises affecting income. This has led to the notion of the double trigger hypothesis: meaning that widespread inability to pay *combined* with low or negative equity that makes selling one's house in the face of financial problems difficult might be more important than ruthless defaults.

Widespread inability to pay stems from two sources. First, the severe recession beginning in 2007 generated substantial income losses across a large swath of households. Second, the sharp rise in nonprime lending during the mid-2000s, including loans with little or no income verification or indeed little or no down payment, likely meant that a substantial fraction of borrowers were financially unstable even at origination. Numerous media anecdotes suggest that ruthless defaults are widespread and the presumed wave of such defaults during the period of the credit crisis has led policy makers to propose ways to deal with ruthless or strategic defaulters.

Financial crises of varying severity have their own distinctive characteristics, and there are also among them a number of significant similarities. One prominent and consistently recurring feature is what the relatively new field of behavioural finance refers to as '*herding behaviour'*, or following the trend, behaviour that although individually rational, produces a group behaviour, that is in a sense irrational.

The analysis of herding constitutes one of the key elements of behavioural finance since only the mistakes of investors at the aggregate level may be reflected in the prices of assets. The behaviour of one individual investor does not affect market prices at all. If investors did not act collectively and did not commit similar psychological mistakes at the same time, their actions would neutralize each other to a great extent, and the market would remain efficient.

This herding behaviour can be seen in action on two levels, amongst mortgage lenders and borrowers, and amongst the investors in mortgage-backed securities (MBS). Each one making their own contribution to explanations of the recent financial crisis in the following manner:

It is known that there needs to be two ingredients for market failure.

1. The existence of low quality products - the flawed incentives of the mortgage securitisation process created mortgage loans of low credit quality with names such as: '*subprime*' or '*non-conforming*' - and

2. Asymmetric information with buyers being less informed than sellers – because not all information about the loans the mortgage brokers had at loan approval was transmitted to the buyers of the MBS. Both these ingredients were present in the credit market crisis.

Once money was lost from the MBS, withdrawal from the market by investors gave rise to a modern version of a '*run on the banks*' in the financial system and the disruption to market function was significant proving debilitating for the economy as the dependence on credit from the market was so great.

After an initial shock from unexpected news when the market temporarily becomes illiquid due to uncertainty, the market normally recovers quickly with trading resuming around new consensus price levels.

However, in this crisis, market recovery did not occur quickly making it obvious that an asymmetric information problem existed. At the time, the existence of asymmetric information was difficult to demonstrate because buyers relied on prices set by credit rating agencies who were presumably fully informed about the approved loans.

Literature testing for herding behaviour and asymmetric information in mortgage markets is thus of mostly recent origin. Empirical work on such asymmetric information is a delicate enterprise, particularly given that what is hidden to the uninformed party is often hidden to the empirical investigator as well. Most of this research has focused on analysing asymmetric information about characteristics of the borrower such as their income prospects. The default risk of mortgages, however, depends both on the value of the housing collateral as well as on the borrower's ability to make interest payments (Deng et al., 2000).

This importance of collateral values in determining mortgage default was particularly visible during the recent crisis, which precipitated many ‘strategic defaults’ in which households that could afford to pay their mortgages chose to default once collateral values fell below the outstanding mortgage balance (Guiso et al., 2010).

Due to the highly illiquid and heterogeneous nature of housing as an asset it is also possible that there is asymmetric information about collateral values in mortgage lending, in particular given the significant resources that lenders spend on appraisals and property inspections to improve their valuation of the housing collateral prior to making a lending decision.

Mian and Sufi (2009) determine that the supply of home credit to non-prime borrowers expanded most rapidly in those areas with the highest rates of house price appreciation that also experienced negative income growth. They largely attribute the increase in mortgage origination in such areas to the rising demand for mortgages to be securitised.

The recent financial bubble that burst began to inflate with mortgage loans made by an array of financial institutions to high risk home buyers, these loans were consolidated into securities packages issued by financial institutions that were then repackaged and sold to other financial institutions and large investors.

The trouble began within a relatively good economic environment in which these financial institutions and large investors search for and seem to concentrate upon trendy, popular assets – the surge in investment in sub-prime mortgage derivatives. The herd grew large with some members seeking to take advantage of the situation as long as house prices are rising everyone prospers. However, when the economy and the housing market starts to decline, going into recession, people lose their jobs. As unemployment rises people start to default on their loans, the number of defaults rises especially on high mortgage payments relative to income as associated with these subprime mortgages, the equity cushion vanishes and there are sharp declines in housing prices. In turn investors begin to lose confidence in these types of assets producing a cascading effect of illiquidity in the market and huge losses for these investors. Eventually a regulatory response occurs.

In general studies of herding behaviour tend to be empirically driven making use of observational data along with evidence gleaned from surveys of mortgage borrowers. While purely ruthless defaults do undoubtedly occur, there is a suggestion that such defaults probably do not account for a large proportion of all defaults in recent years, see Stroebelz (2011).

Herding in asset markets can be seen as reflecting peoples' response to problems of uncertainty and limited information. *Herd behaviour*, or herding, is generally not considered to be a bias, but more of a behavioural tendency. However it may be hypothesised to influence the mortgage decision making processes, both for borrowers and advisors.

In an uncertain world, housing investors will not only look at their best judgements of the likely future path of an assets value, they will also look to the valuations of other investors in deciding what to do either because the crowd may have superior information or because there is security in numbers. This is because if they realise that they do not have complete information, then it will be sensible to look at what others are doing before deciding how to value their own housing assets and decide whether or not to buy or sell now or later.

To summarise, herding behaviour is *fragile*, in that it may break easily with the arrival of some trigger (i.e. new information). It is *idiosyncratic*, in that random events combine with the choices of the first few participants to determine the type of behaviour on which individuals herd. Herding is also *dynamic* in that a small change in one parameter can quickly change the systems behaviour, i.e. if house price increases do not stimulate supply sufficiently to counteract the price rise.

Herding comes in *waves* in that delay is followed by a sudden simultaneous action and it is *path dependent*, as the outcome depends on the order of activities.

Herding effects tend generally to magnify the instabilities initially generated and the level of financial default. '*Housing frenzies*' emerge as a result - if the level of repossessions is high, the market responds to this and the herd follows a pessimistic path, leading to falls in housing transactions and magnifying the initial effects of default. Conversely, if house prices are rising, the herd follows an optimistic path amplifying the impact of house price appreciation. So, upward frenzies will be created when prices are rising, downward frenzies emerging when repossessions are rising meaning that overall, the timing of housing decisions will not be evenly spread over time. Due to its inherent non-linearity, herding is thought to be better modelled by non-linear dynamic systems.

In this thesis, an assessment of borrower herding behaviour - of how borrower behaviour conforms to the ruthlessness hypothesised in option theoretic models - is not directly addressed. Qualitatively an assessment of the effects of transactions costs is provided however a more formal assessment of the suboptimal behaviour of borrowers - as attributed to Transaction Costs - is also not directly addressed by this thesis. Rather, these effects are taken to be an unobservable influence, forming part of the general heterogeneity of borrower behaviour examined in aggregate in the empirical data analysis of Chapter 8 but could however form the basis of further future investigation.

3.2.5 Concluding Remarks

Before describing the main characteristics from the perspective of financial modelling it may be helpful to quickly review the most relevant elements in the valuation of a mortgage contract.

The valuation of the mortgage asset depends upon two main factors; these are the structure of its cash flows and the economic environment in which it is traded. Mortgages are not securities in themselves like stocks and shares since they are not (usually) traded. However, a mortgage can be treated as a derivative security, it may be viewed as a fixed income instrument combined with the embedded options to prepay and default.

In a contingent claims valuation framework the elements that are used to characterise the economic environment are the most relevant sources of risk inherent in the asset under study. In the case of mortgages these are the interest rate term structure and the dynamics of the house price; the values of both are subject to economic uncertainty and are thus modelled by suitable stochastic processes, identified in detail in the next section.

Under normal circumstances in a world of uncertainty in which economic participants react to risk in different ways it is necessary to take into account investor preferences. However, one of the main achievements of the contingent claims approach is that these preferences need not be established, their role being relatively marginal in the valuation of derivative assets.

The cash flow structure of borrower payments is determined by the mortgage contract rate which is itself dependent upon the market interest rate along with certain other features of the mortgage contract such as its Term and any teaser discounts applied in the first year. The contract rate is fixed at mortgage origination so as to not allow the possibility of an arbitrage situation to exist.

The other contractual features of almost all types of mortgage are the two embedded options forming the possible sources of early termination. The borrower is given the option to terminate the loan prior to maturity date, through either prepayment or through default.

The value of mortgage debt to the lender is reduced by the option to prepay or to default since the borrower may prepay or default, not both as discussed in Appendix A. Further, when default occurs, the borrower not only surrenders the house but also terminates the prepayment option and similarly, by prepaying the borrower foregoes the default option. Therefore default and prepayment options should not be valued independently, but valued jointly as they constitute competing risks of early termination.

The primary determinant of mortgage prepayment is the term structure of interest rates. The two primary determinants of mortgage default are the borrower's current equity position in the mortgaged property - loan to value LTV ratio - and the size of the borrower's mortgage payment obligation, relative to his disposable income - payment to income PTI ratio.

In computing the value of the mortgage to the borrower it is also assumed that markets are efficient and that borrowers act rationally and optimally exercise their default and prepayment options. In reality as the literature review of chapter two highlights many borrowers may sub-optimally exercise the default or prepayment options due to personal or non-financial reasons. Consequently, the market value of a mortgage will reflect both optimal and suboptimal default and prepayment from various borrowers.

A great deal of the literature review of chapter two deals with research predominantly undertaken in the USA and refers specifically to the US mortgage market. However, there is no reason conceptually to support the idea that borrower behaviour in the UK mortgage market is any different from that in the US so that the major determinants of termination in the UK should not be any different from those in the US.

Chapter 4

An Options Pricing Model of Mortgage Valuation

4.1 Introduction

This section presents the founding framework for the valuation model derived in this thesis. This includes both the definition and mathematical description of the various elements that comprise the framework. Namely, of the two-state variables governing the valuation equation, the valuation equation itself, the identification of the components of the mortgage contract and the structure and value of the mortgage payments to be made by the borrowers.

4.1.1 Economic Environment

For residential mortgages the model assumes that in an uncertain economic environment the factors affecting the returns from a mortgage can be summarised by two state variables: the entire term structure of interest rates and the value of the underlying house prices. The term structure of interest rates is assumed to be generated from the stochastic process describing the spot interest rate, $r(t)$. This follows a mean reverting square root diffusion process (Cox, Ingersoll and Ross, 1985b). The value of the underlying house prices, $H(t)$, is also uncertain and is assumed to follow a log-normal diffusion process (Merton, 1973). The set up comes from Kau, Keenan and Kim, (1994). The stochastic processes are depicted in the equations Eqn (3.1) and Eqn (3.3).

The *Term Structure* process is written as*:*

$$dr = \gamma(\theta - r)dt + \sigma_r \sqrt{r} dz_r \qquad \ldots \quad \ldots \text{ Eqn (3.1)}$$

where

$\gamma \equiv$ The rate of adjustment in the mean reverting process

$\theta \equiv$ The long term mean of the short-term interest rate, *r(t)*

(Steady state spot rate)

$\sigma_r \equiv$ Instantaneous standard deviation of the interest rate disturbance

$z_r \equiv$ Standard Wiener process

This process indicates that the interest rates revert at rate γ toward a steady-state value θ, but that they are being constantly disturbed by stochastic events, as represented by the Weiner process z_r. The Volatility of these interest rate disturbances is described by the parameter $\sigma_r \sqrt{r}$. The equation assures that if $\gamma \geq 0$ and $r \geq 0$ initially, then subsequent negative interest rates are precluded, for this process zero is a natural reflecting barrier. Notice also that the equation lacks an interest-rate risk premium, that is, it is already risk neutral.

This follows from the Local Expectations Hypothesis (LEH). The LEH is that the expected change in the value of a pure discount bond is simply equal to the instantaneous interest rate:

$$\frac{E[dP(t,T)]}{P(t,T)} = r(t)dt \qquad \ldots \quad \ldots \qquad \text{Eqn (3.2)}$$

where *P(t, T)* is the value of a pure discount bond at time *t* paying one dollar at time *T*. It essentially serves to prevent there being any risk premia in the term structure. A discussion of the LEH can be found in Cox Ingersoll and Ross (1981).

It is assumed that the interest-rate risk premium has been absorbed by the interest rate parameters $\theta\, and\, \gamma$ (Cox, Ingersoll and Ross, 1979).

The house price process is originally adapted from the stock process in the tradition of Black and Scholes (1973) and Merton (1973). One parameter that must be included when modelling the house value is the service flow, or implicit rent, which acts as a 'dividend'. The return to owning the house therefore comes in two parts as price appreciation and as a service flow from using the house over time. This service flow is assumed to be $\phi H dt$ proportional to the value of the house and is subtracted from the assets total expected return since the holder of an option on the house has no claim to this service flow. Note that this process has an absorbing barrier at zero, meaning that if H(t) ever becomes zero, it remains at zero thereafter.

The *House Price* process is written as:

$$\frac{dH}{H} = (\alpha - \phi)dt + \sigma_H dz_H \qquad \ldots \quad \ldots \qquad \text{Eqn (3.3)}$$

This is an Ito process (geometric Brownian motion) where

$\alpha \equiv$ The instantaneous average rate of house price appreciation

$\phi \equiv$ 'dividend type' per unit service flow provided by the house

$\sigma_H \equiv$ Instantaneous standard deviation of the house price

$z_H \equiv$ Standard Wiener process

Here ϕ is the constant service or rental flow from the house, assumed to be proportional to the value of the house, whereas α is the required expected return on the housing asset, so that $(\alpha - \phi)$ represents the expected rate of appreciation in the house price. It is assumed that houses are traded at no risk premium, which allows α to be replaced by *r*.

Once again, there are stochastic disturbances in this trend, whose magnitudes are described by the volatility parameter σ_H. These disturbances may be correlated with fluctuations in the term structure, with such a correlation being captured in the coefficient ρ, where

$$dz_r dz_H = \rho dt \qquad \ldots \quad \ldots \qquad \text{Eqn (3.4)}$$

where the standard Wiener processes are such that E[dz]=0 and E[dz^2]=dt.

$\rho(r, H, t)$ is the correlation between the disturbances to the term structure and house price. Whether the correlation coefficient should be positive or negative is dependent upon whether nominal or real forces dominate Kau and Keenan (1980, 1981) for the real case and Kau and Keenan (1983) for the nominal case. Titman and Torous (1989) estimate the correlation coefficient to be between –0.1 and –0.3.

4.2 The Structure of the Solution

4.2.1 Preliminaries

Mortgages are considered as being derivative assets. Their prices are assumed to depend on the evolution of the term structure of interest rates and house prices. Once the house price and term structure are determined, the value of the mortgage contract is set through a process of arbitrage inference.

All other factors that may exert an influence will only be taken into consideration through the market price of risk associated with each state variable.

The primary features of a mortgage contract are its amortising payments and the opportunity to terminate the contract through either prepayment or default. Prepayment and default are endogenously determined by the borrowers rational behaviour to minimise the value or cost of the mortgage contract, *V(H,r,t)*, so that 'optimal' termination is assumed.

Now, the cost of a mortgage is an expected present value calculation, and hence of the general representation provided by the Feynman-Kac theorem:

$$V(H(t),r(t),t) = \hat{E}[e^{-\int_t^\tau r(v)dv} V(H(\tau),r(\tau),\tau] \qquad \dots \quad \dots \qquad \text{Eqn (3.5)}$$

for $t \leq \tau$, integral between t and τ, where the terminal value $V(H(\tau),r(\tau),\tau)$ of the mortgage contract at expiration date τ, is specified by the terms of the contract and the caret on *E* denotes that the stochastic processes must be 'risk adjusted'.

By standard arguments in finance, Smith (1976), based around the perfect capital market assumption together with the Local Expectations Hypothesis, the present value of the mortgage contract *V(H, r, t)*, at time *t* is of the form of Eqn(3.5)

The assumed absence of arbitrage possibilities among mortgage asset values requires a linear pricing rule (Cox, Ingersoll and Ross,1985a, b) which is equivalent to adjusting the stochastic processes in terms of the so called market prices of risk, λ_i, one for each source of uncertainty, *i*, and then taking expectations in risk-neutral fashion with respect to the now risk-adjusted processes:

$$dr = [\gamma(\theta - r) - \lambda_r \sigma_r \sqrt{r}]dt + \sigma_r \sqrt{r} dz_r \qquad \dots \quad \dots \qquad \text{Eqn (3.6)}$$

and

$$dH = [(\alpha - \phi) - \lambda_H \sigma_H]Hdt + \sigma_H H dz_H \qquad \dots \quad \dots \qquad \text{Eqn (3.7)}$$

Now assuming that the property is a continuously traded asset, then since the economy is to appear risk-neutral after adjustment, it must be that the risk-adjusted expected return to the house is simply the risk free rate (this is as a consequence of the Girsanov theorem in that there exists a 'risk-adjusted measure Q, say, under which actualised prices of perfectly tradable assets are Martingales. This implies that under Q, the risk adjusted expected returns of these assets are simply the risk free rate, *r*.) so $[(\alpha - \phi) - \lambda_H \sigma_H]H / H + \phi = r$ that on rearranging gives

$$[(\alpha - \phi) - \lambda_H \sigma_H]H = (r - \phi)H \qquad \ldots \qquad \ldots \qquad \text{Eqn (3.8)}$$

On the other hand, the effect of the Local Expectations Hypothesis is to assure that the interest rate risk factor is null, that is $\lambda_r = 0$. Thus in order for the solution to be correct we must substitute in for the true stochastic processes (1) and (2) the so called *risk-adjusted* processes, which are in our case:

$$dr = \gamma(\theta - r)dt + \sigma_r \sqrt{r}\, dz_r \qquad \ldots \qquad \ldots \qquad \text{Eqn (3.9)}$$

and

$$\frac{dH}{H} = (r - \phi)dt + \sigma_H dz_H \qquad \ldots \qquad \ldots \qquad \text{Eqn (3.10)}$$

Notice that in the course of transforming the house process to its risk adjusted form, all reference to α the required rate of return on the house disappears. Thus, the values of the mortgage and of the default option embedded in the mortgage are independent of α.

Under this framework, with the stochastic processes represented in Eqn (3.9) and Eqn (3.10) it is known that the partial differential equation (PDE) for the valuation of any asset *V(r, H, t)* whose value is a function only of the two mentioned state variables and time takes the form

$$\frac{1}{2}H^2\sigma_H^2\frac{\partial^2 V}{\partial H^2}+\rho H\sqrt{r}\sigma_H\sigma_r\frac{\partial^2 V}{\partial H\,\partial r}+\frac{1}{2}r\sigma_r^2\frac{\partial^2 V}{\partial r^2}+\gamma(\theta-r)\frac{\partial V}{\partial r}+(r-\phi)H\frac{\partial V}{\partial H}+\frac{\partial V}{\partial t}-rV=0$$

… … Eqn (3.11)

The value of a particular mortgage contract can then be obtained by imposing the appropriate terminal and boundary conditions on the partial differential equation.

4.2.2 Derivation of Model Dynamics

In this section the PDE that describes how the option values evolve over time is discussed and its limitations are highlighted. The derivation of the asset valuation PDE itself is relegated to Appendix D and follows from standard arguments in finance and in particular upon the '*no arbitrage*' condition. As we can see in the derivation of the PDE this allows us to assert that a riskless portfolio will have a return equal to the riskless rate, *r*.

Now certainly the stock market operates in a way that is indeed much closer to the assumptions of arbitrage theory that the Housing Market. Consider the case of the prepayment option where the first problem with the theoretical situation stems from the uniqueness of the housing asset itself. If a trader attempted to build an arbitrage on a prepayment option he cannot own the specific house associated with the mortgage because it is already owned by the borrower! He can only attempt to build a riskless portfolio on a similar housing asset adding a risk of imperfect replication. Houses cannot be substituted one for another as two stocks can. A second problem arises if the strategy implies the sale of a house. If the trader does not own the house it could be difficult to take a short position on a house.

And finally property prices are not quoted as stocks are. The actual price of a house is only revealed at the time of processing the sale transaction, thus establishing price uncertainty until a real transaction is made. This leads to the notion of an 'error' of pricing, associated with house price appraisal, for a contingent claim based upon housing assets (with consequent implications for the type of UK equity release loans called *'lifetime mortgages'*).

The physical housing asset and its service flow can - however infrequently - be bought and sold and so if for the purposes of modelling we consider it to be traded in a perfect and frictionless market (where traders can buy and sell, lend or borrow in order to make money from the small inefficiencies within the mortgage contracts) then mortgage asset values can be determined by the solution to the PDE. However, options embedded in mortgage contracts have prices that depend more upon competition between lenders than upon hypothetical arbitrages, so this brings into question the validity of the arbitrage pricing method in this instance - since the financial characteristics of housing assets are not in good agreement with the requirements of arbitrage theory.

In conclusion, the valuation PDE is a very powerful result fundamentally assuming some very strong conditions on the underlying asset securities under study.

In a sufficiently well developed, mature, liquid and frictionless market (in other words in a "perfect" financial market) the no-arbitrage assumption is acceptable.

It is this assumption that is essential to the model as it allows us to assert that a riskless portfolio must give a return equal to the riskless rate r, for producing finally the asset valuation equation. In this section we have acknowledged however, that the situation in the housing market is far from this. The imperfections involved are not secondary difficulties which may be resolved by merely tweaking the assumptions a little, as can be done with stochastic volatility models.

The imperfections here bring into question the model as a whole, and ultimately, the accuracy and validity of the final PDE. The modelling is not therefore sufficiently linked to housing market particulars, and so one must bear in mind that the direct application of arbitrage theory, originally built for "perfect" financial products such as stocks, can produce potentially misleading valuations.

Note that time-varying assets relying on interest rate products are sufficiently standardised, negotiated and liquid that we may reasonably consider these to be perfectly tradable assets in a frictionless market; they can be bought or sold without any limitation so that building an arbitrage based strategy on lending or borrowing money is not a problem provide one makes clear the specific model for the risk neutral process for r. Since there is already a very well established theory for the pricing of interest rate products their validity is not brought into question in this section.

4.2.3 Restrictions on the valuation PDE

For this section the problem of Fixed Rate Mortgage (FRM) Valuation is explored and a detailed explanation of the model is provided.

Since all financial claims contingent on a given asset share the same general valuation equation, the characteristics that distinguish one such asset from another are entirely embodied in the terminal and boundary conditions needed to close the model. Here the properties of the standard amortising fixed-rate mortgage are laid out and in doing this the terminal conditions of its constituents are specified. To assist the discussion and to formalise the model some notation is introduced:

Formulae:

L: The loan amount (the amount of the debt at the origination of the loan)

c: The fixed yearly contract interest rate (coupon rate for repayment mortgage)

n: The Term of the loan in months (the life of the mortgage in months)

i: The Payment date in months $1 \leq i \leq n+1$

$\tau(i)$: The calendar time of the i^{th} month i.e. $\tau(i) = i / 12$ the i^{th} payment date

MP: Fixed monthly mortgage payment

$OB(i)$: The Unpaid principal (Outstanding Balance) after the i^{th} payment date

$TD(\tau)$: The Unpaid principal plus accrued Interest for $\tau(i) < \tau \leq \tau\,(i+1)$

$A(r, \tau)$: Value at time τ of the promised mortgage payments from i to n

$D(H, r, \tau)$: Value at time τ of the default option when the next mortgage payment is due at time $\tau(i)$

$P(H, r, \tau)$: Value at time τ of the prepayment option when the next mortgage payment is due at time $\tau(i)$

$V(H, r, \tau)$: Value at time τ of the mortgage contract when the next mortgage payment is due at time $\tau(i)$

Value of Monthly Payments: The present value of the monthly payment *MP* is given by the standard annuity formula

$$L = MP \frac{[1-(1+\frac{c}{12})^{-n}]}{\frac{c}{12}}$$ which on rearranging gives

$$MP = \frac{L(\frac{c}{12})(1+\frac{c}{12})^{n}}{(1+\frac{c}{12})^{n}-1}$$ gives the formula for the value of the mortgage payments.

. . . . . . Eqn (3.12)

Value of Principal Balance: The unpaid principal outstanding at time *i*, is determined the future value of the outstanding debt in the final period is equal to the future value of all the payments. That is $L(1+\frac{c}{12})^{n}$ is to be met by $MP(1+\frac{c}{12})^{n}$ so that the outstanding balance the borrower still has to pay at time *i,* immediately after the i^{th} monthly payment has been made will be expressed by:

$$OB(i) = L(1+\frac{c}{12})^i - MP(1+\frac{c}{12})^i[\frac{(1-(1+\frac{c}{12})^{-i})}{\frac{c}{12}}]$$ substituting for MP:

$$OB(i) = L(1+\frac{c}{12})^i \left[1 - \frac{(1+\frac{c}{12})^n[1-(1+\frac{c}{12})^{-i}}{(1+\frac{c}{12})^n - 1}\right]$$ and simplifying gives:

$$OB(i) = L\left[\frac{(1+\frac{c}{12})^n - (1+\frac{c}{12})^i}{(1+\frac{c}{12})^n - 1}\right]$$ . . . . . . Eqn (3.13)

giving the formula for the outstanding balance *OB(i)* after the i^{th} monthly payment has been made.

Value of Unpaid Principal plus Accrued Interest:

The unpaid principal plus accrued interest $TD(\tau)$ for $\tau(i) < \tau \leq \tau\ (i+1)$ is given by

$TD(\tau) = OB(i)\ [\ 1 + c(t - \tau(i))]$ . . . . . . Eqn (3.14)

Joint Prepayment and Default Option Value:

In order to calculate the value of the mortgage for a borrower it is necessary to account for the present value of the future payments promised to the lender and the value of the options implicit in the mortgage contract. At any time, the value of the joint option to terminate the mortgage prior to maturity *J(H, r, t)* will be given by *J(H, r, t) = P(H, r, t) + D(H, r, t)*

Consequently representing the value of the remaining payments by *A(r, t)* the value of the mortgage to the borrower will be

$$V(H, r, t) = A(r, t) - J(H, r, t) = A(r, t) - P(H, r, t) - D(H, r, t)$$

. . . . . . Eqn (3.15)

4.2.4 PDE and Terminal Conditions

The valuation problem consists of valuing several different assets simultaneously. Each asset has its own payment-date conditions for each month of the mortgage. To determine the value of each asset at mortgage origination, the valuation PDE must be solved backwards in time. The valuation procedure differs between the maturity of the loan and the other payment dates.

At payment dates a distinction has to be made between the value of the asset before and immediately after each payment.

$F^{-}(H, r, t) =$ Value of asset immediately before payment is made

$F^{+}(H, r, t) =$ Value of asset immediately after payment is made.

At any time step during the valuation of the mortgage the value of the Asset *F(H ,r, t)* must be calculated for all house price values and all interest rate values. Here unless otherwise stated *F(H, ,r, t),* represents any of the following mortgage components:

V(*H, r, t*) - the value of the mortgage to the Borrower

D(*H ,r, t*) – the value of the option to Default

P(*H, r, t*) – the value of the option to Prepay

Value of Promised Mortgage Payments

Valuation of the promised future payments, involves only the term structure of interest rates. At Maturity $(t = T)$: The value of the scheduled monthly payment is $A^{-}(r,t) = MP$ the value of the last monthly payment.

Value of Mortgage

At Maturity $(t = T)$: The borrower may either pay the required amount *MP* or default. To the lender its value immediately before the payment at maturity is the minimum of *MP* and the house value so $V^{-}(H,r,t) = \min(MP, H)$.

Immediately after payment of the final instalment the mortgage is fully amortised thus the values of the options vanish.

Value of Prepayment Option

At Maturity $(t = T)$: The option to Prepay at maturity is by definition meaningless so its value is zero so $P^{-}(H,r,t) = 0$

Value of Default Option

At Maturity $(t = T)$: The default option will be worthless if the value of the house is more than the value of the final payment. Otherwise it will be equal to the difference between the two $D^{-}(H,r,t) = \max(0,(MP - H))$

At other Payment Dates $(1 < t < T - 1)$

Given here are the values of all the mortgage components for other payment dates - that is at the end of each month 1, 2, 3 and so on.

Moving backwards in time as each monthly payment date is reached the borrowers debt reduces by the amount MP, as $A^{+}(r,t)$ becomes $A^{-}(r,t)$ Hence:

$$A^{-}(r,t) = A^{+}(r,t) + MP\,. \qquad \text{Eqn (3.16)}$$

Adding the natural boundary condition to the terminal conditions stated, it is possible to obtain the value of *A(r, t)* at the origination of the mortgage loan A(r(0), 0).

The value of the Mortgage to the lender is the lesser of *MP* plus its value after the payment and the house price itself so

$$V^{-}(H,r,t) = \min((V^{+}(H,r,t) + MP), H) \qquad \text{Eqn (3.17)}$$

Default occurs if the value of the house is less than the sum of the monthly payment plus the value of the mortgage immediately after the payment is made. In this situation the option components take on the following values

In default: $D^{-}(H,r,t) = A^{-}(r,t) - H$ if $V^{-}(H,r,t) = H$

No default: $D^{-}(H,r,t) = D^{+}(H,r,t)$ if $V^{-}(H,r,t) = V^{+}(H,r,t) + MP$

If Default occurs then prepayment option cannot be exercised giving

In default: $P^{-}(H,r,t) = 0$ if $V^{-}(H,r,t) = H$

No default: $P^{-}(H,r,t) = P^{+}(H,r,t)$ if $V^{-}(H,r,t) = V^{+}(H,r,t) + MP$

Insurance

Considering now the insurance perspective. If the borrower chooses to exercise the option to default, the lender will lose future payments that would have been received had default not occurred. The lender may take advantage of insurance to cover a fraction of this loss. This Insurance is termed a Mortgage Indemnity Guarantee (MIG). Now, the value of a mortgage that has a MIG associated with it is not the same for the lender as for the borrower.

Mortgage Indemnity Insurance. The MIG is an insurance asset that is not part of the *'mortgage package'* of embedded assets taken into account by a financially rational borrower. Only the lender benefits from the guarantee provided by this insurance policy.

The MIG, $I(H,r,t)$, only adds to the lender's position in the contract. However, its value is dependent on the mortgages expected performance.

Let the value of the mortgage to the lender be $V_L(H, r, t)$, then the value of the contract for the lender will be given by $V_L(H, r, t) = V(H, r, t) + I(H, r, t)$ its value is thus only known once the value of the contract to the borrower $V(H, r, t)$ is known. Borrower behaviour plays a significant role in the definition of mortgage value and as the work of this thesis is structured to derive the value of the mortgage contract to the borrower then throughout this work references to 'mortgage value' are understood to mean ' mortgage value to the borrower'.

To model the MIG: The insurer agrees paying a fraction λ, of the total loss, TD (t) – H suffered by the lender, but only subject to a maximum indemnity or cap of Λ. It may be assumed that this cap Λ = 0.2 times the original value of the house and that λ =0.8 based upon Azevedo-Pereira et al. (2000, 2002, 2003).

The Conditions at Maturity are

In default:

The MIG has value $I^{-}(H,r,t) = \min(\lambda[MP - H], \Lambda)$ if $V^{-}(H,r,t) = H$

No default:

The MIG has value $I^{-}(H,r,t) = 0$ if $V^{-}(H,r,t) = MP$

At other payment dates

In default:

The MIG has value $I^{-}(H,r,t) = \min(\lambda[TD^{-}(t) - H], \Lambda)$ if $V^{-}(H,r,t) = H$

No default:

The MIG has value $I^{-}(H,r,t) = I^{+}(H,r,t)$ if $V^{-}(H,r,t) = V^{+}(H,r,t) + MP$

Co-insurance. Coinsurance is the fraction of the potential loss not covered by the MIG it includes the loss above the cap. Its value is useful to the insurer and lender but also to any third party who may have interest in selling cover for this source of contract risk. At each payment date the value of the coinsurance is the difference between the potential loss and the cover provided by the MIG.

Letting *CI (H, r, t)* represent the value of this coinsurance then at the final payment date, the Conditions at Maturity are:

In default: The Coinsurance has value

$$CI^{-}(H,r,t) = \max((1-\lambda)[MP - H], (MP - H) - \Lambda) \qquad \text{if } V^{-}(H,r,t) = H$$

No default: The Coinsurance has value

$$CI^{-}(H,r,t) = 0 \qquad \text{if } V^{-}(H,r,t) = MP$$

At other payment dates:

In default: The Coinsurance has value

$$CI^{-}(H,r,t) = \max((1-\lambda)[TD^{-}(t) - H], [TD^{-}(t) - H] - \Lambda) \quad \text{if } V^{-}(H,r,t) = H$$

No default: The Coinsurance has value

$$CI^{-}(H,r,t) = CI^{+}(H,r,t) \qquad \text{if } V^{-}(H,r,t) = V^{+}(H,r,t) + MP$$

4.2.5 PDE and Boundary Conditions

The mortgage valuation model requires Boundary Conditions to close the problem. Now, the natural boundary conditions are those that specify the value of the contract when the state variables take on extreme values that is (H=0 and $H \to \infty$) along with (r=0 and $r \to \infty$) and the Terminal Condition is that at contract expiry.

However, with each monthly mortgage payment we have a ‘partial’ termination condition. This is dealt with by treating the mortgage as compound contract, so that the solution of the contract over a later month provides the terminal condition to start the solution procedure over the previous month, thus the solution proceeds iteratively, backwards in time.

Condition along H=0, (House value is zero):

Under this condition the absolute value of the Mortgage cannot be less than the price of the house.

A borrower's rational behaviour is to default so, the prepayment option is worthless, the value of the mortgage is equal to the value of the house and the value of the default option is just the value of the remaining mortgage payments since *D* = *A*-*V*-*P*.

$P(0,r,t) = 0$; $V(0,r,t) = H = 0$; $D(0,r,t) = A(r, t)$

If the borrower chooses to exercise the option to default, the lender will lose future payments that would have been received had the default not occurred.

In this case the value of any insurance related products the lender may have (Mortgage Indemnity Guarantee, *I(H,r,t)* and Coinsurance, *CI(H,r,t)*) are given by the degenerate version of the valuation PDE

$$\frac{1}{2}\sigma_r^2 r^2 \frac{\partial^2 I}{\partial r^2} + \gamma(\theta - r)\frac{\partial I}{\partial r} + \frac{\partial I}{\partial t} - rI = 0 \text{ and}$$

$$\frac{1}{2}\sigma_r^2 r^2 \frac{\partial^2 CI}{\partial r^2} + \gamma(\theta - r)\frac{\partial CI}{\partial r} + \frac{\partial CI}{\partial t} - rCI = 0$$

Condition along *r*=0, (Interest rate is zero):

The value of all the other components will be given by the solution of the degenerate form of the valuation PDE when *r* is considered to be null. That is

$$\frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 F}{\partial H^2} - \phi H \frac{\partial F}{\partial H} + \gamma\theta \frac{\partial F}{\partial r} + \frac{\partial F}{\partial t} = 0$$

Condition at *r*=0, *H*=0, (Both House value and Interest rate is zero):

Here the value of any Asset *F(0,0,t)* will be given by $F(0,0,t) = F(0, \gamma\theta s, t + s)$

The value of the remaining future payments A(r, t) is dependent only upon the term structure of interest rates and time. In the situation where the interest rate *r*, is null there will be no discounting. Given the CIR mean reverting square root process, the value of the interest rate in the next time step, *s*, is certain to be $\gamma\theta s$ and so the boundary condition for the value of future payments, A, at any moment in time is given by

$$A(0,t) = A(\gamma\theta s, t + s)$$

Condition at $r=0$ as $H \to \infty$,

Similarly the value of the assets will be given by

$$\lim_{H\to\infty} F(H,0,t) = \lim_{H\to\infty} F(H,\gamma\theta s, t+s)$$

Condition along $H \to \infty$, (House value is very large):

Here default cannot have any value, so $\lim_{H\to\infty} D(H,r,t) = 0$ but this is not the case for prepayment.

In this case the value of this option is a function of the relationship between the amount of outstanding debt and the value of the mortgage. The value of the mortgage is constant as *H* tends to infinity implying that $\frac{\partial V}{\partial H} \to 0$

The degenerate form of the valuation PDE that gives the mortgage value is then

$$\frac{1}{2}\sigma_r^2 r^2 \frac{\partial^2 V}{\partial r^2} + \gamma(\theta - r)\frac{\partial V}{\partial r} + \frac{\partial V}{\partial t} - rV = 0$$

Since the value of the mortgage is the difference between the value of the remaining future payments and the borrower's joint option to terminate, then prepayment option value at this extreme will be

$$\lim_{H\to\infty} P(H,r,t) = A(r,t) - \lim_{H\to\infty} V(H,r,t)$$

In the situation where default is not possible the insurance coverage is worthless

$$\lim_{H\to\infty} I(H,r,t) = \lim_{H\to\infty} CI(H,r,t) = 0$$

Condition along $r \to \infty$ (Interest rate is very large):

In the limit of infinite interest rate, any expected future payment is worthless. So in accordance with Azevedo-Pereira et al. (2000,2002,2003)

$$\lim_{r\to\infty} A(r,t) \to 0$$

When $r \to \infty$ any future payment is worthless consequently

$\lim_{r\to\infty} V(H,r,t) = \lim_{r\to\infty} D(H,r,t) = \lim_{r\to\infty} P(H,r,t) = 0$ and

$\lim_{r\to\infty} I(H,r,t) = \lim_{r\to\infty} CI(H,r,t) = 0$

Condition as $r \to \infty$ $H \to \infty$, (Both House value and Interest rate have extreme values):

As seen previously assets involving payments in the future are of no value. Consequently, when $r \to \infty$ the equations giving the value of the assets there, continue to hold.

$\lim_{r\to\infty} V(H,r,t) = \lim_{r\to\infty} D(H,r,t) = \lim_{r\to\infty} P(H,r,t) = 0$ and

$\lim_{r\to\infty} I(H,r,t) = \lim_{r\to\infty} CI(H,r,t) = 0$

4.2.6 Treatment of the Free Boundary

The Default Boundary: The option to default is a compound European option, since it can only be exercised on the payment date in a given month. Default cannot occur if the option to prepay is exercised. Thus, default is only rational outside the prepayment region and so the default boundary is fully described by the payment-date conditions.

The Prepayment Boundary: The option to prepay is American in type, since prepayment could occur at any time during the lifetime of the mortgage contract. This produces a free boundary, which must be applied in the appropriate position that is by those combinations of *H* and *r* that correspond to points where prepayment first takes place. On one side of this free boundary it is financially optimal for a borrower to prepay on the other side it is not. Prepayment boundary conditions are obtained by a consideration of the following.

Given the rational behaviour of borrowers, then the total debt *TD*, at each moment in time must be at least as high as the value of the mortgage *V*.

$$V(H, r, t) \leq TD(t) \qquad \text{Eqn (3.18)}$$

The prepayment boundary is then given by the region where $V(H, r, t) = TD(t)$ that is a *'value matching'* condition is observed. In addition to this there is the requirement of a *'smooth pasting'* condition, Merton (1973), that both function *V* and *TD* meet tangentially at the boundary.

This means that their slopes should match at the boundary which, then translates with respect to the state variables as

$$\frac{\partial V}{\partial H} = \frac{\partial TD}{\partial H} = 0 \text{ And } \frac{\partial V}{\partial r} = \frac{\partial TD}{\partial r} = 0 \text{ when } V = TD.$$

If prepayment takes place, then the default option becomes immediately worthless. The same happens with the insurance related assets since there is no possibility of default.

Default, Insurance and Coinsurance make sense only outside the prepayment region. Since there is a need to observe smoothness in the solutions for these along the prepayment boundary then

$$\frac{\partial D}{\partial H} = \frac{\partial I}{\partial H} = \frac{\partial CI}{\partial H} = 0 \text{ and } \frac{\partial D}{\partial r} = \frac{\partial I}{\partial r} = \frac{\partial CI}{\partial r} = 0$$

If the decision to prepay the mortgage is made, then the borrower is liable to pay the lender an amount that is referred to as the Total Debt *TD*.

This may include a penalty for the early termination of the mortgage contract and the way in which the penalty is calculated does not appear to be standardised across mortgages. It will be modelled here as a proportion of the outstanding balance plus accrued interest at the moment of termination.

In this situation the amount of the total debt in the event of prepayment will be given by

$$TD(t) = (1+\pi)(1+c(t-\tau(i)))OB(i) \qquad \text{Eqn (3.19)}$$

where π represents the early termination penalty, $\tau(i)$ is the time of the i^{th} month payment date, c is the fixed coupon rate of the mortgage and *OB(i)* is the outstanding balance after each payment date.

Free Boundaries are difficult to treat numerically and there are two approaches to deal with them. Firstly there are boundary-tracking methods these are discussed by Crank, (1984). Secondly there is the use of transformations that are capable of reducing the problem to a fixed boundary one, from which the free boundary can be inferred, these are reviewed by Wilmott et al. (2001b).

The solution here is to adopt this latter approach drawing upon the work of Berger, Ciment and Rogers (1975) where the problem is converted into a non-linear PDE with a fixed boundary.

4.2.7 Equilibrium Condition

The terms of the mortgage contract need to be agreed by the two parties involved so as to avoid arbitrage. This means that the original contract must be in a state of equilibrium. This will be so if the value of the mortgage to the lender is the same as the amount lent to the borrower.

The equilibrium condition can be expressed in a general way as follows,

$$V(H(0), r(0), t(0), d, \pi) + I(H(0), r(0), t(0), d, \pi) = (1-\xi)L \qquad \text{Eqn (3.20)}$$

Here the *lender's position* in the contract is *V*=*A*-*D*-*P*, that is, the scheduled payments, *A* minus the sum of the borrowers joint options to terminate the mortgage *D* (by default) and *P* (by prepayment) plus any Insurance *I*, the lender may hold against the borrower defaulting on promised payments.

The *borrower's position* is the amount lent, *L*, which will be some percentage of the initial house value minus an arrangement fee ξ that is charged as a percentage of the loan amount. This fee is non-reimbursable and acts as a deterrent to the early termination of the contract along with the early termination penalty π also imposed at the beginning of the contract as a percentage of the loan amount.

For an equilibrium situation to be reached, an equilibrium contract rate, *d*, needs to be found that is capable of balancing Eqn(3.20) above at origination of the contract. The equilibrium contract rate, *d*, which balances the equilibrium constraint, can be found easily using an iterative process based upon Newton's Method in line with Gerald and Wheatley (1994) and Press et al. (1992). In numerical analysis, this method is a root-finding algorithm (based upon a linear approximation of the function) that uses a succession of roots of tangent lines that better approximate the root of a function *f*.

Let *f*(*d*) be a function of the contract rate only, where d is the contract rate used to balance the equilibrium constraint, then *f*(*d*) is given by

$$f(d) = V(H(0), r(0), t(0), d, \pi) + I(H(0), r(0), t(0), d, \pi) - (1-\xi)L \qquad \text{Eqn (3.21)}$$

that must be equal to zero to satisfy the constraint.

An initial estimate for *d* is made call this $d_{0.}$ The values of the mortgage components involved in the equilibrium constraint are calculated with this initial estimate as starting point. Next, a tolerance to which the absolute value of *f*(*d*) must be less than is specified; once the point at which *f*(*d*) is less than this tolerance is reached, the iterative procedure terminates. It is only at this point that *d* is the equilibrium contract rate.

For the iteration process an estimate is also required for the incremental change in $d_{0.}$ Call this increment Δ_0 (specified) then the next potential equilibrium setting d is given by $d_1 = d_0 + \Delta_0$. Using this detail it then possible to calculate $f(d_1)$ and then check if its absolute value is less than the tolerance.

If the absolute value of $f(d_1)$ and any further $f(d_i)$ is greater than the tolerance then a new increment for the change in d will be calculated by the following recurrence formula

$$\Delta_{i+1} = -\left(\frac{\Delta_i f(d_i)}{f(d_i) - f(d_{i-1})} \right) \text{ where } i \geq 1 \qquad \text{Eqn (3.22)}$$

Newton's method is widely used and is chosen here because in the near neighbourhood of a root, it is rapidly convergent.

4.2.8 Model Assumptions, Weaknesses and Validity: A Review.

Before concluding this present chapter it would be useful to review the assumptions upon which this finalised model is to be presented. The basic assumption of the model that is common to most of the option pricing theory is that we do not know tomorrow's values of the underlying asset prices due to the complexity of the financial market. The past history of the asset price is available and can be examined, but cannot be used to forecast the next move the asset will make.

Almost all option-pricing models are founded on one simple model for asset price movement involving parameters derived from historical or market data. It is assumed that the asset prices must move randomly, because of the *'efficient market hypothesis'*, which can be briefly described by the two statements:

(a) The past history is fully reflected in the present price, which does not hold any further information.

(b) Markets respond immediately to any new information about an asset.

We also assume the following no-arbitrage or *arbitrage free argument* is valid. That is the return on the portfolio and the return on a riskless bank account is the same. The '*no-arbitrage argument*' implies that in a financial market there are no opportunities to make instantaneous risk-free profits. Hence, the modelling of asset prices can be interpreted as the modelling of the arrival of new information which affects the asset price. With these assumptions, changes in the asset price define a Markov process.

In addition to the no-arbitrage assumption the mortgage value is also assumed to satisfy another arbitrage-free condition. The contract rate must be so that the *equilibrium condition* holds.

In the modelling framework of this thesis, it is assumed there are two sources of uncertainty: the default risk and term structure or the interest rate risk. The default risk is mainly tied to the random fluctuations of the property price, which are traded assets. On the other hand, the interest rate risk is tied to random perturbations of financial markets and the behaviour of the economy as a whole. The term structure risk is thus much more complex, since interest rates are not directly traded. For these reasons, two specific stochastic processes are chosen that are assumed to characterize these two risks.

Interest Rate

In the literature many of the most popular interest rate models focus only on the short rate $r(t)$. In this thesis a short rate model is considered, where the future behaviour of interest is stochastic and only one interest rate the short (spot) rate $r(t)$ is specified to predict the future behaviour of interest rates., This has the desired characteristics of:

(a) $r(t)$ should be positive (desired of all interest rates).

(b) $r(t)$ should be mean-reverting (also called autoregressive), $r(t)$ cannot drift off to plus or minus infinity or to zero, but must eventually be pulled back to some long-term target.

(c) Simple formulae for the prices of interest rate based contracts should be obtained (a matter of computational convenience rather than of economic principle.).

The Cox-Ingersoll-Ross CIR model is one that satisfies on all of these assumptions consequently the CIR model is the interest rate model adopted by most of the literature on mortgage valuation and is adopted in this thesis as well.

Property Price

In the literature the property price is almost always modelled by a geometric Brownian motion, ensuring the inclusion of the service flow or implicit rent that is taken to be proportional to the value of the house. It is assumed that houses are traded at no risk premium and this assumption is also adopted in this thesis. The stochastic disturbances of the term structure and property price are assumed correlated.

These two processes for the house price H(t) and the short interest rate r(t), are assumed to capture all the sources of uncertainty in the standard approach to mortgage valuation, therefore H(t) and *r*(t) are adopted as state variables. Our model thus assumes that the factors affecting the returns from a mortgage - the mortgage value - can be summarised by and will depend on them as well as on time t.

Time Reversal

Before moving on, a small change of variable, which is just a change in the direction of time, from forwards to backwards looking is required. This modification is not essential, but it simplifies the numerical schemes that are used to approximate the solution to the mortgage value.

Option Exercise

The model assumes Competing Risks of early termination and that termination events are determined endogenously. The value of the mortgage is the same as scheduled payments, minus the value of the borrower's options to terminate the mortgage.

Before a borrower continues making the scheduled monthly payments, he must make sure that his mortgage has a value that is smaller than the total debt (house value - plus any transactions costs) otherwise it is not advantageous for him and it would be better for him to prepay the mortgage by exercising the prepayment American Call option. This is mathematically expressed as a free boundary condition, where the prepayment option is not exercised if the following inequality holds: Value<Total Debt for all times to expiry. Similarly for a default event to occur the value of the mortgage exceeds the value of the underlying property asset.

The model also assumes that the mortgage is treated as a derivative asset and borrowers act ruthlessly in their rights to terminate the mortgage so acting rationally in an optimising way (maximising wealth / minimising costs).

The temporal conditions, together with the boundary conditions and the free boundary condition are all that is necessary to find the value of a mortgage and its related components.

Modelling Approach

A contingent claims framework for valuing the mortgage and the associated embedded options is used in this thesis. The advantage of this methodology lies in providing a template for the borrower, lender and insurer to compare mortgage terms, including the fairness of contract rates, fees and any insurance premiums required, when borrowing more than 75% of the house value at origination.

In this framework there are two approaches to derive pricing formulas from a short rate model, one is called the Martingale approach, and the other is the Partial Differential Equation (PDE) approach. The Martingale approach uses the theory of martingales to establish prices, the PDE approach is a general approach very similar to the option pricing approach as developed by Black-Scholes.

Although the martingale approach is generally thought to be more powerful and intuitive than the option-theoretic PDE approach, the latter models still provide a useful tool for the development of numerical methods and provide useful insights into the future values and behaviour of mortgage values and they are widely used in the type of modelling followed in this thesis.

In the case of mortgage valuation models and due to the complexity of mortgage contracts, there is really no hope for a simple valuation formula, a closed-form solution does not actually exist. Mortgages are, among the most complex contracts ever devised, so to value them, most models resort to numerical methods to find an approximate value. This is indeed preferred, since it is preferable to have an accurate value of the right model, rather than an analytical formula of the wrong one. Numerical methods for the valuation of option-theoretic mortgage models typically involve finite difference approximations, backward methods, which rely on the fact that the value of a mortgage is known at the time of maturity and, given the economic environment, one can work backward in time to find the value of a mortgage at a previous instant of time for all possible values of interest rate and house prices through numerical simulations.

In option-theoretic mortgage models, the right to prepay is typically regarded as a call option - an '*American*' option, starting at mortgage origination and having the same expiration time as the mortgage itself, while the right to default is thought of as a put option – actually a sequence of linked monthly '*European*' options, that start right after a monthly payment is made and have expiration at the next payment date.

A mortgage model must, at the very least, consider the possibility of early termination by means of either default or prepayment; reasons for defaulting or prepayment are complex. Financial reasons play an important role in the case of default - applying equally well to all individual borrowers, while personal reasons are the main influence in the decision to prepay, and an attempt to incorporate such complexity is made in this thesis.

Given the conventional terms and the diffusion processes, for the simulation studies we assume the parties have common expectations for some diffusion parameters, including expected future rate and price volatilities enabling the setting of initial parameters.

Also we assume that borrowers preferences need not be established, that behaviour of borrowers in the UK is identical to those in the USA and that borrowers are ‘economic actors’ regarding default and prepayment options, and those factors such as mobility, divorce, unemployment, and transaction costs are deterministic and can be ignored in building this model, along with moral hazard problems between the insurer and the lender.

Weaknesses

Standard mortgage models like the one presented in this thesis are good approximations to real world situations, especially when the house price H increases and the interest rates are somehow ‘high’, as was the case before the housing market crisis that started in 2007.

However, when the house price H drops, the standard model would not give very good approximations, as several situations that are not considered in these models may suddenly occur and significantly affect the value of a mortgage.

It would be no longer reasonable to assume that real world randomness is completely summarized by the chosen stochastic process equations.

Thus to account for unforeseen scenarios that the standard mortgage modelling approach does not consider, this thesis incorporates the following.

1. A change to the stochastic process governing property prices by making them mean reverting. This changes the effect played by the house service flow.

2. A *discount factor* that will, in general, decrease the value of the mortgage depending on the most current economic conditions available. This *discount factor* is carefully chosen so as to include up-to-date information on several aspects of the economy.

Therefore, mortgage values calculated with these approaches will be more realistic and accurate.

The structural model only considers one type of default/prepayment trigger: financial and assumes perfect capital markets; that is, there are no transaction costs, no taxes, and no informational asymmetries. The model assumes that under conditions of limited liability borrower behaviour is rational, ruthless in application and not suboptimal.

The model thus has a number of other limitations briefly noted below. Due to the highly illiquid and heterogeneous nature of housing as an asset it is possible that there is *asymmetric information* regarding borrower income as previously identified but also about property values in mortgage lending, in particular given the significant resources that lenders spend on property appraisals and property inspections to improve their valuation of the housing collateral prior to making a lending decision.

The model does not consider the effects of *herd behaviour*, or herding, discussed earlier in the thesis as it is generally not considered to be a bias, but more of a behavioural tendency. However it may be hypothesised to influence the mortgage decision making processes, both for borrowers and advisors. In general, studies of herding behaviour and the effects of asymmetric information tend to be empirically driven making use of observational data along with evidence gleaned from surveys of mortgage borrowers.

Our model has further limitations. In particular, it is technically very difficult to introduce other factors into this framework to take into account other important aspects of the default / prepayment decision, such as the risk of income loss and exogenous reasons for residential mobility and refinancing.

The former would require the introduction of a third stochastic variable which would make the solution of the model extremely complex.

Extensions of this modelling framework can be made by modifying the assumptions governing some aspect of the valuation problem, and some of these possibilities are considered in this thesis.

In the light of the analysis presented, structural models seem most suitable to provide the basis for general equity-credit models. However, they need to be extended theoretically. Theoretically, they need to include in some form the possibility of default due to liquidity shortages and high financing costs.

Model Validity:

Despite the above limitations we believe the model is indeed a valid one. The models assumptions are limiting only to a minor extent, we argue this for the following reason. The model uses option theory and this establishes that the appropriate statistical and economic theory has been used in the development of the model and provides formulae for equity, debt and default probabilities.

We ensured the validity of data input and the statistical assumptions through consideration of historic data aspects this ensured that the data used (initial simulation parameters) for results are representative of their application to the mortgage market and consistent with economic reasoning.

Theoretically, then the structural model does indeed provide a satisfactory framework for jointly modelling equity and credit risks.

4.3 Conclusion

This chapter presents a theoretical framework for the valuation of UK mortgages and Mortgage Indemnity Guarantees. At the level of the repayment mortgage, the terminal conditions imposed take into account the specific nature of the early prepayment penalties included in most UK mortgages. At the level of the insurance valuation, the terminal conditions account for features included in some UK policies like the cap and the definition of the guarantee as a proportion of the loss. In addition the simultaneous valuation of coinsurance, the potential loss not covered by the MIG but covered by the lender is allowed. No apparent closed-form solutions are available for the resulting PDE Eqn (3.11).

In the presence of the special difficulties created by the free boundary imposed by the American Option to prepay the mortgage loan and the need to also cope with the compound European Option to default at payment dates, a numerical solution of the PDE Eqn (3.11) is required. The next chapter presents a numerical procedure to solve the valuation PDE Eqn (3.11) using a finite difference method.

Chapter 5

Numerical Solution Method

5.1 Introduction

Most problems arising in financial mathematics cannot be solved analytically. Instead use is made of numerical techniques to obtain their solution. There are several popular methods in the literature, which are used to approximate the value of derivative securities like mortgages. These methods include Monte Carlo method, Lattice Methods (binomial and trinomial trees) and finite difference methods. This study will involve only the latter method. Section 5.2 presents a summary of each these numerical methods.

The remainder of this chapter is organised as follows. Section 5.3 describes the structure of the numerical solution. Section 5.4 demonstrates how the transformed valuation PDE of the previous section is represented by an explicit finite difference approximation. Section 5.5 concludes by discussing the issues of stability and convergence of the applied explicit finite difference approximation.

5.2 Brief Review of Techniques

Monte Carlo Approach. The Monte Carlo approach is a forward method, in that a solution starts from the initiation of the option at time t=0. Random sample paths are then generated (for instance using a normal distribution) according to which stochastic process is used to model the underlying asset. Sample paths are discounted at a specified interest rate to determine the implied option value.

Boyle (1977) was the first to utilise Monte Carlo simulation in the solution of option valuation problems. As far as mortgage valuation is concerned Monte Carlo simulation is effectively ruled out. This is a result of the methods main limitation; that the simulated paths are generated forwards in time, making it difficult to determine when it is optimal to exercise and so value options which have early exercise features. This is indeed the case with mortgages in determining the optimal time for a borrower to prepay.

Lattice Approaches. The Lattice approach is backward method that can readily handle early exercise features. Lattice or tree methods were developed independently by Rendlemann and Bartter (1979) and Cox et al. (1979). A simple Illustration of the lattice approach is given in an Appendix H

The fundamental reasoning behind this approach is that at each discrete moment in time, the asset price can either move up or down to a new level in a binomial lattice or in the case of trinomial lattice to a third level somewhere in between. The value of the option is known at expiry *T*, i.e. its payoff, and so this can be used to evaluate the price at $T - \Delta t$. This is performed recursively in order to determine the value of the asset at t=0.

Early exercise features are no problem here as the valuation process is performed backwards in time enabling a comparison to be made with the market price of the option, if early exercise is taken. Despite the fact that lattices are more intuitive and easier to implement the lattice approach has the limitation that the tree structure is not very flexible.

It is difficult for example to align its nodes with important asset prices such as a barrier or the exercise price, accordingly finite difference methods seem to be the most widely used.

Finite Difference Approach. Most problems that arise in mathematical finance are often best solved directly by finite difference techniques. These numerical techniques are ideally suited to optimal-stopping problems. Brennan and Schwartz (1977) recognised the utility of this method to price American options using an explicit finite difference scheme. This is simply because the transformation from a PDE to the difference equation(s) is easier when the lattice's grid is nice and regular. Solution via a backward method makes it simple to track the free boundary that arises when determining the optimality of American option exercise.

Finite difference techniques are extremely flexible and can allow for the inclusion of complex path-dependent option features, such as those found in barrier options. Enhanced or improved variations of these methods exist and are used today with better accuracy and wider application.

The framework for option pricing is built around the Black Scholes (1973) and Merton (1973) equation. This is a backward parabolic equation that may be solved analytically in simple cases. In many problems that are adaptations and generalisations of this equation (mortgages being treated as derivative securities being no exception) however, it cannot, and so must be most efficiently solved using finite difference methods.

Finite difference methods allow an intuitive feel for the problem and how the solution may be obtained. The underlying problem is converted from one that exists in a continuous domain to one that can be described in a discrete finite domain. The derivatives of the PDE Eqn (3.11) are discretised, being replaced by linear difference equations and the state space becomes a regular grid/mesh on which the problem is defined.

5.3 The Structure of the Numerical Solution

In this section, the PDE Eqn (3.11) is first transformed from a backwards parabolic PDE to a forward equation by reversing the time dimension. The PDE is solved numerically by transforming the variables, namely, converting the coefficients of the PDE to constants, yielding a linear form to which the boundary conditions are then more accurately and easily applied. The initial conditions are also changed to forward in time.

The numerical solution of the valuation PDE Eqn (3.11) requires a number of technical difficulties to be overcome. The first one relates to its infinite domain.

In the interest rate dimension (r-space): Arbitrage arguments require that the time value of money cannot be negative, this means that the natural domain of the PDE is the space $[0,\infty)$.

In the House Price dimension (H-space): In a housing market without any major inefficiency and also without transaction costs and finite supply of housing, it would follow that house prices should also be non-negative. The natural domain of the PDE is the space $[0,\infty)$.

Unfortunately, infinite boundary conditions are difficult to handle by numerical methods and so this situation requires the use of a transformation of the original PDE. The chosen transformation should be capable of minimising the problems that arise as a result of this particular feature.

The transformation chosen is that of Stanton (1995). Instead of working with *r* and *H* directly, we use a transformation of variables to *x* and *y*, say.

In the interest rate dimension the transformation is $y = \dfrac{1}{1+\varphi r}$ for some constant $\varphi > 0$ so that when $r = 0, y = 1$ and as $r \to \infty, y \to 0$.

The inverse transformation is $r = \dfrac{1-y}{\varphi y}$

In the House Price dimension the same transformation yields $x = \dfrac{1}{1+\omega H}$ for some constant $\omega > 0$ so that when $H = 0, x = 1$ and as $H \to \infty, x \to 0$.

The inverse transformation is $H = \dfrac{1-x}{\omega x}$

Under these transformations the infinite region $[0,\infty)\times[0,\infty)$ is mapped onto the unit square $[0,1]\times[0,1]$

The time variable also needs to be transformed as Wilmott (2001b) mentions that parabolic PDE's should preferably be solved in the forward sense. Hence the transformation used is $\tau = T - t$ the inverse transformation being $t = T - \tau$

We may now express the original function $V(h(x),r(y),t(\tau)) = W(x, y, \tau)$ in terms of these new variables as follows:

For the *First derivatives* we obtain

$\frac{\partial V}{\partial r} = \frac{\partial W}{\partial y}\frac{\partial y}{\partial r}$ by chain rule and $\frac{\partial y}{\partial r} = \frac{(1+\varphi r).0 - 1.\varphi}{(1+\varphi r)^2} = -\varphi y^2$ by quotient rule, and $\frac{\partial V}{\partial H} = \frac{\partial W}{\partial x}\frac{\partial x}{\partial H}$ and $\frac{\partial x}{\partial H} = -\omega x^2$

likewise for the *Second derivatives*

$\frac{\partial^2 V}{\partial r^2} = \frac{\partial W}{\partial y}\frac{\partial^2 y}{\partial r^2} + \frac{\partial y}{\partial r}\frac{\partial^2 W}{\partial y^2}\frac{\partial y}{\partial r} = \frac{\partial W}{\partial y}\frac{\partial^2 y}{\partial r^2} + \frac{\partial^2 W}{\partial y^2}\left(\frac{\partial y}{\partial r}\right)^2$ by the product and chain rules

$\frac{\partial^2 y}{\partial r^2} = \frac{(1+\varphi r)^2.0 - 1.(-\varphi.2(1+\varphi r).\varphi)}{(1+\varphi r)^4} = 2\varphi^2 y^3.$

Similarly $\frac{\partial^2 V}{\partial H^2} = \frac{\partial W}{\partial x}\frac{\partial^2 x}{\partial H^2} + \frac{\partial^2 W}{\partial x^2}\left(\frac{\partial x}{\partial H}\right)^2$ and $\frac{\partial^2 x}{\partial H^2} = 2\omega^2 x^3$ finally for the mixed derivative

$$\frac{\partial^2 V}{\partial r \partial H} = \frac{\partial y}{\partial r}\frac{\partial}{\partial H}\left(\frac{\partial W}{\partial y}\right) = \frac{\partial y}{\partial r}\left(\frac{\partial^2 W}{\partial y \partial x}\right)\frac{\partial x}{\partial H}$$

For the Time derivative

$$\frac{\partial V}{\partial t} = \frac{\partial W}{\partial \tau}\frac{\partial \tau}{\partial t} \text{ And } \frac{\partial \tau}{\partial t} = -1$$

Substituting these elements into the valuation PDE and simplifying gives the formulation of the transformed PDE as follows:

$$\frac{1}{2}\sigma_H^2 \left(H(x)\right)^2 \omega^2 x^4 \frac{\partial^2 W}{\partial x^2} + \frac{1}{2}\sigma_r^2 \left(r(y)\right)\varphi^2 y^4 \frac{\partial^2 W}{\partial y^2} +$$
$$\rho\sigma_H \sigma_r H(x)\sqrt{r(y)}\varphi\omega x^2 y^2 \frac{\partial^2 W}{\partial x \partial y} +$$
$$\left[\sigma_H^2 (H(x))^2 \omega^2 x^3 - (r(y) - \phi)\omega x^2 H(x)\right]\frac{\partial W}{\partial x} +$$
$$\left[\sigma_r^2 r(y)\varphi^2 y^3 - \gamma(\theta - r(y))\varphi y^2\right]\frac{\partial W}{\partial y} - \frac{\partial W}{\partial \tau} - r(y)W = 0$$

... ... Eqn (4.1)

Now under the transformations $y = \dfrac{1}{1+\varphi r}$ and $x = \dfrac{1}{1+\omega H}$ the number of points on a given *x(y)* grid that corresponds to *H(r)* values is inversely proportional to the value of $\omega(\varphi)$. Therefore the smaller the values of (φ) and (ω) the more points there will be on a given grid.

To determine the values of (φ) that are used to give plausible values of *r* that will be seen to occur in the market during the 25yr time horizon of a mortgage, we require the midpoint of the grid to be a reasonable value for *r*, say 10%

Then $y = \dfrac{1}{2}, r = 0.10 \Rightarrow \varphi = 10$

Similarly, $x = \dfrac{1}{2}$ and at par *H* = 1 so $\omega = 1$ (remains Fixed) so that we can simply scale up by the actual value of the House.

The second as discussed in the next section relates to the technique by which the continuous partial differential equations are discretised so that they can be represented and solved on a computer.

In the finite difference method, the partial derivatives of the governing equations are discretised directly by replacing the derivatives with difference equations and approximating the solution of the PDE by a set of algebraic equations and solving on a discrete set of points.

5.4 Discrete Representation - Derivative Approximations

In order to develop approximations for derivatives in terms of these discrete points *Taylor Series Expansions* are used see Appendix E for details. By using these expansions the transformed state space in the x and y dimensions is discretised using central difference approximations for the first and second derivatives and the time dimension is discretised using a forward difference approximation.

Upon substituting these derivative approximations into the transformed valuation PDE and rearranging, the problem of solving this PDE reduces to solving a set of simultaneous linear equations as outlined in the section below.

5.4.1 Representing the valuation PDE by Finite Difference Approximations.

At any time step during the valuation of the mortgage, the value of the asset must be calculated for all house-price and interest-rate values. Moving to the discrete representation of the problem, this means that this value must be found all for *i* and *j*.

The two dimensional state space is to be subdivided into a number of discrete intervals the complete set of which is called a grid.

The Interest Rate (*y*-) dimension will be represented in the Interval [0,1] as being partitioned into J intervals of length, k, such that $y_j = jk$ and $Jk = 1$.

Similarly, the House Price (*x*-) dimension will be represented in the Interval [0,1] as being partitioned into I intervals of length, h, so that $x_i = ih$ and $Ih = 1$.

Finally the Time dimension will be represented in the interval [0,*T*] as being subdivided into N intervals of length, s, so that $Ns = T$ and $t_n = ns$

The value of the asset *W(x ,y, τ)* will then be approximated by $U_{i,j}^n$ at the node or lattice point *P(i, j)*. Utilising this notation, on the interior nodes of the lattice, those corresponding to points where 0 < *x* < 1 and 0 < *y* < 1, the space derivatives can be expressed in terms of central difference approximations.

The time derivative is represented by a forward difference approximation (see Wilmott et al., 2001). This leads to the partial derivatives of the transformed Valuation PDE being written in terms of the finite difference approximations as follows:

$$\frac{\partial W}{\partial x} \cong \frac{U_{i+1,j}^n - U_{i-1,j}^n}{2h}\,;\ \frac{\partial W}{\partial y} \cong \frac{U_{i,j+1}^n - U_{i,j-1}^n}{2k}\,;\ \frac{\partial W}{\partial t} \cong \frac{U_{i,j}^{n+1} - U_{i,j}^n}{s}$$

$$\frac{\partial^2 W}{\partial x \partial y} \cong \frac{U_{i+1,j+1}^n - U_{i+1,j-1}^n - U_{i-1,j+1}^n + U_{i-1,j-1}^n}{4hk}$$

$$\frac{\partial^2 W}{\partial x^2} \cong \frac{U_{i+1,j}^n - 2U_{i,j}^n + U_{i-1,j}^n}{h^2}\ ;\ \frac{\partial^2 W}{\partial y^2} \cong \frac{U_{i,j+1}^n - 2U_{i,j}^n + U_{i,j-1}^n}{k^2}$$

This then leads the valuation PDE to be represented as follows:

Let $a(x,y,t) = \frac{1}{2}\sigma_H^2 (H(x))^2 \omega^2 x^4$

Let $b(x,y,t) = \frac{1}{2}\sigma_r^2 r(y) \varphi^2 y^4$

Let $c(x,y,t) = \rho\sigma_H\sigma_r H(x)\sqrt{r(y)}\varphi\omega x^2 y^2$

Let $d(x,y,t) = \sigma_H^2 (H(x))^2 \omega^2 x^3 - (r(y)-\phi)\omega x^2 H(x)$

Let $e(x,y,t) = \sigma_r^2 r(y)\varphi^2 y^3 - \gamma(\theta - r(y))\varphi y^2$

then

$$a(x,y,t)\left[\frac{U_{i+1,j}^n - 2U_{i,j}^n + U_{i-1,j}^n}{h^2}\right] + b(x,y,t)\left[\frac{U_{i,j+1}^n - 2U_{i,j}^n + U_{i,j-1}^n}{k^2}\right] +$$

$$c(x,y,t)\left[\frac{U_{i+1,j+1}^n - U_{i+1,j-1}^n - U_{i-1,j+1}^n + U_{i-1,j-1}^n}{4hk}\right] + d(x,y,t)\left[\frac{U_{i+1,j}^n - U_{i-1,j}^n}{2h}\right] +$$

$$e(x,y,t)\left[\frac{U_{i,j+1}^n - U_{i,j-1}^n}{2k}\right] - \left[\frac{U_{i,j}^{n+1} - U_{i,j}^n}{s}\right] - r(y)U_{i,j}^n = 0$$

Multiplying throughout by s and then rearranging to make $U_{i,j}^{n+1}$ the subject gives

$$U_{i,j}^{n+1} = \left(1 - sr(y) - \frac{2as}{h^2} - \frac{2bs}{k^2}\right)U_{i,j}^n + \left(\frac{bs}{k^2} + \frac{es}{2k}\right)U_{i,j+1}^n + \left(\frac{bs}{k^2} - \frac{es}{2k}\right)U_{i,j-1}^n +$$

$$\left(\frac{as}{h^2} + \frac{ds}{2h}\right)U_{i+1,j}^n + \left(\frac{as}{h^2} - \frac{ds}{2h}\right)U_{i-1,j}^n +$$

$$\frac{cs}{4hk}\left(U_{i+1,j+1}^n - U_{i+1,j-1}^n - U_{i-1,j+1}^n + U_{i-1,j-1}^n\right)$$

. . . . Eqn (4.2)

This shows a relationship that gives that gives the asset value at a certain time step (*n*+1) as a function of its own value at the previous time step (*n*).

A consideration of the coefficients indicates that not only are they functions of the model parameters but also of the values of the transformed space variables;

so they are themselves variable. Unfortunately this makes the chosen scheme prone to generating stability and hence convergence problems.

The applied finite-difference approximation of the transformed equation leads to the leading-order error terms, $O(\delta H^2, \delta r^2, \delta t,)$ - The error in the explicit method for the problem set up in this thesis - which are responsible for time-step restrictions and numerical instability as discussed in the next section.

5.5 Stability and Convergence Issues

There are two main approaches to implementing finite difference methods in the solution of Eqn (4.2): explicit and implicit schemes. Explicit and implicit finite difference algorithms can adopt a multitude of different representations of the time and space differentials.

According to the formulation adopted in Wilmott (2001b), both schemes represent the space differentials by symmetric differences. The major difference between the two alternatives is the way in which time differentials are represented.

If a forward difference is used to represent the time differential, the resulting difference equation involves only one grid point at the next time level - the moment in time that is treated subsequently to the current one by the algorithm. This method is called the *Explicit* finite difference approximation. At any moment in time, the value of the function at the next time step is represented in terms of its own values in several space points at the current time step.

In the *Implicit* method if a backward difference is used to represent the time differential, the resulting difference equation involves more than one grid point at the next time level. In this case there is a need to solve a set of simultaneous equations, giving the structure of the solution added complexity compared to that of the Explicit approximation method. This does not mean that the method should be avoided in the solution of PDE's, since in practice one finds that there is no one '*best'* method for approximating difference formulae.

There are advantages and disadvantages associated with each approach. A choice in the face of any situation is required. It is generally known that Implicit methods have better stability characteristics but Explicit methods solutions are easier to obtain. The choice of method used in a given situation is the one that is thought to perform the best in terms of a number of attributes. These are as follows.

Accuracy – measures the deviation between the numerical solution and the exact solution of the differential equation.

Stability – indicates the assurance that the deviation between the solutions at a fixed point in time is always bounded. That is the rounding errors that arise as a consequence of the finite precision arithmetic of computers are not magnified at each iteration. The numerical solution should itself remain uniformly bounded.

Consistency – gives the assurance that the difference equation solution is convergent to the correct solution, as the finite difference mesh is refined. That is, truncation errors should vanish as the mesh sizes involved and the time step tend to zero.

A numerical method is said to be *Convergent* if the difference between the theoretical solutions of the differential and difference equations at a fixed point in time tends to zero uniformly as the dimension of the space steps in the mesh tends to zero. That is the finite-difference solution approaches the true solution to the partial differential equation.

5.5.1 Stability and Convergence of the Numerical Procedure.

For the purposes of the work in this thesis to keep error within acceptable bounds it is required to assure that all the U^n coefficients are positive (Morton and Mayers, 1994). In order to achieve this goal, there are two possibilities. The first consists of a reduction of the spatial step size. This substantively reduces the viability of the scheme, so it is proposed to fix the spatial step size at $h=k=0.02$, In other words, the numerical solution is obtained from 50 × 50 grid; the second consists in a change in the representation of the first derivative terms.

The coefficients of the first derivative terms change sign across the grid so according to Morton and Mayers (1994, p48) a solution consists in using a forward difference representation when the coefficient is positive and a backward difference representation when it is negative, this idea being known as *'upwind differencing'*.

The reason for the use of upwind differencing for the state variables first derivatives is that the volatility of the state variables is usually quite small.

As the transformed equation is two-dimensional, there are four possible combinations of first derivative coefficient signs:

(i) d > 0, e > 0 implying a Forward difference representation for both x and y first derivatives

(ii) d > 0 e < 0 implying a Forward difference representation for x and a Backward difference representation of y first derivatives

(iii) $d < 0$ $e > 0$ implying a Backward difference representation for x and a Forward difference representation of y first derivatives

(iv) $d < 0$ $e < 0$ implying a Backward difference representation for both x and y first derivatives

Note also that the coefficients of the second derivative terms are always positive in the transformed equation.

No formal error theory for this problem appears to exist and consequently studies that utilise this type of finite difference technique do not draw close attention to this subject.

As relatively little is known in this respect, (Henrici, 1963 and Schurz, 2001) then in order to gain some understanding and insight it is necessary to consider rather specialised problems and so the commonly observed approach is therefore to consider the general heat flow equation as an approximation to the problem in hand in order to provide insight into the conditions required for convergence to be reached, since from the Lax Equivalence Theorem, it is known that the stability and consistency of a finite difference scheme are both necessary and sufficient conditions to ensure convergence. (Smith, 1978, p72)

With this in mind, a reasonable insight into the conditions which must be fulfilled in order to obtain convergence in our case can be gleaned from a consideration of this approach as adopted in Morton and Myers (1994).

To apply this general approach to our problem, note that the convergence of the explicit finite difference method depends upon the size of the time step, size of the space step and size of the derivative coefficients.

For our two state variable valuation problem, recall that the general expression for the explicit difference equation Eqn (4.2) may be written as in the following:

$$a(x,y,t)\left[\frac{U^n_{i+1,j}-2U^n_{i,j}+U^n_{i-1,j}}{h^2}\right]+b(x,y,t)\left[\frac{U^n_{i,j+1}-2U^n_{i,j}+U^n_{i,j-1}}{k^2}\right]+$$

$$c(x,y,t)\left[\frac{U^n_{i+1,j+1}-U^n_{i+1,j-1}-U^n_{i-1,j+1}+U^n_{i-1,j-1}}{4hk}\right]+d(x,y,t)\left[\frac{U^n_{i+1,j}-U^n_{i-1,j}}{2h}\right]+$$

$$e(x,y,t)\left[\frac{U^n_{i,j+1}-U^n_{i,j-1}}{2k}\right]-\left[\frac{U^{n+1}_{i,j}-U^n_{i,j}}{s}\right]+f(x,y,t)U^n_{i,j}=O(\delta H^2,\delta r^2,\delta t)$$

Where a, b, c, d, e, f are all the coefficients in the PDE previously defined.

In order for the numerical solution of this PDE to converge and be stable, the following conditions must be satisfied (Wilmott, 2001b)

(i) $f \leq 0$

(ii) $\frac{|d|}{a}\delta H+\frac{|e|}{b}\delta r\leq 2$

(iii) $a\frac{\delta t}{\delta H^2}+b\frac{\delta t}{\delta r^2}\leq\frac{1}{2}$

Now in general financial problems, there is invariably a negative *f*; (often it is simply –*r* where *r* is the risk free rate) and as this is the situation in our case, then condition (i) is satisfied.

The second constraint on the space steps can be avoided by the use of the upwind differencing method. As this method is being used here in this case, then condition (ii) is satisfied

The third constraint however puts a serious limitation on the size of the time step. The application of this constraint to our problem implies that for convergence and stability as *h*, *k*, *s* tend to zero the following condition needs to be satisfied:

$$\frac{1}{2}\sigma_H^2 (H(x))^2 \omega^2 x^4 \frac{s}{h^2} + \frac{1}{2}\sigma_r^2 r(y)\varphi^2 y^4 \frac{s}{k^2} \le \frac{1}{2}$$

where *h* = the grid spacing in the x-dimension;

k = the grid spacing in the y-dimension;

s = the grid spacing in the time dimension.

And since

h = *k*, this then becomes

$$\frac{\left(\sigma_H^2 (H(x))^2 \omega^2 x^4 + \sigma_r^2 r(y)\varphi^2 y^4\right)s}{h^2} \le 1 \qquad \ldots \qquad \ldots \qquad \text{Eqn (4.3)}$$

Notice that the value of the left hand side changes with both the parameters that are used to characterise the volatility of the economic environment and the parameters involved in the transformation process. Consequently we need to be assured that the step size, *s*, for the time variable will be able to provide a guarantee of validity, for each point in the grid along with each set of economic environment parameters.

Given the spatial grid to achieve this with a space step *h*=*k*=0.02 (50 steps) and following Kau, Keenan, Muller and Epperson (1993) and Azevedo-Pereira et al (2000, 2002, 2003), to achieve this with a margin of security, for all circumstances considered in the present work we can assure the satisfaction of the stability and convergence criteria with 66 time steps per month corresponding to s ≈ 0.00126. (That is 66s = 1/12)

Using the base parameter values and points of the grid that would typify the situations that could potentially cause the most problems, the following table presents results of the calculations for testing whether the time step constraint holds good.

Table 1: Stability and Convergence Conditions

Position in Grid				Scale Factors		Economic Environ-ment		α	β	$\alpha\frac{s}{h^2}+\beta\frac{s}{k^2}$
x	$H(x)$	y	$r(y)$	ω	ψ	σ_H	σ_r			
0.98	0.02	0.02	4.92	1	10	0.1	0.1	1.9E-06	3.9E-07	7.3E-06
0.98	0.02	0.98	0.002	1	10	0.1	0.1	1.9E-06	9.4E-04	2.97E-03
0.5	1.0	0.5	0.1	1	10	0.1	0.1	3.1E-04	3.1E-03	0.0109
0.02	49.0	0.02	4.92	1	10	0.1	0.1	1.9E-06	3.9E-07	7.3E-06
0.02	49.0	0.98	0.002	1	10	0.1	0.1	1.9E-06	9.4E-04	2.97E-03

Note: The space step is h=k=0.02 and the time step is s=0.001262

The last column of the table above demonstrates that condition (iii) in this case is indeed satisfied

The solution will be based upon a grid where the interval [0,1] is partitioned into 50 steps. To discretise each of the space dimensions in this way requires h = k = 0.02. In other words the numerical solution is obtained from a 50 by 50 state space grid where h = k = 0.02.

To discretise the time dimension in months for the duration T (= 25 years) of the mortgage means that we need to partition the interval [0,T] as follows, since we have $\frac{1}{T}[0,1]$ years or $\frac{12}{T}\left[0,\frac{1}{12}\right]$ months then we need to divide this interval into N steps of size s, sufficient to satisfy the certain stability and convergence criteria previously discussed so that s = 0.00126.

In conclusion, the computer code written to solve the problem with the given spatial step sizes and time step size allows for the choice of the approximation to be dictated automatically conditional on the sign of the first derivative in each dimension thus ensuring that all alternatives are considered.

In this thesis the issues of the propagation of error and of ensuring convergence to local optima have been considered and found acceptable by investigation of the two kinds of error in the numerical method. First is the discretisation error: by determining that the conditions identified on page 138 are satisfied, we ensure that it is a reasonable expectation for the truncation errors to vanish as the step sizes involved tend to zero.

Second is the local round-off error of the computing machine. These round-off operation errors of the computer are affected by how numbers are represented in the machine itself – their precision, the order in which arithmetic operations are arranged and other details of the developed programming code,

However, whilst not indicating quantitatively what magnitude of errors are actually to arise rather it provides for a reasonable expectation that the accumulated round off errors are bounded.

A detailed investigation the propagation of the errors involved is beyond the scope of this thesis, however for practical purposes, as the Lax Equivalence Theorem indicates stability + consistency = convergence.

This was investigated numerically, by comparing the results computed on a series of successively refined grids, where it was found that the proposed solution grid and time step was better in terms of computational speed and memory requirements along with the absence of system crashes through floating point overflow.

Chapter 6

Numerical Simulation Results

6.1 Introduction

In contrast with securities that are traded on Stock Exchanges, mortgages are far away from being a standardised product.

The characteristics of contracts vary, between lending institutions. There are thus a large number of possibilities that could have been tested. Given the computational load and the level of computer processing ability available, a considerable amount time is required to generate complete sets of numerical results.

Even when generated results are subject to a certain degree of post processing in order to provide a reasonable picture of the evolution of the valuation functions inherent in the model under different simulated conditions. Therefore it has been necessary to restrict the number of cases presented to just a few.

A decision was made to consider fixed rate repayment type mortgages as these largely represent Mortgage Market. The repayment mortgages considered differ in respect to the nature their negative incentives to early termination. For instance, by the inclusion of an arrangement fee a contract is made close to the structure of American mortgages. However the level of the arrangement fee is slight less than the level of *'points'* that seems to be commonly charged in the US. UK Mortgages also have early termination penalties that affect exercise of borrower options and as a result equilibrium contract rates.

In addition there is also the insurance coverage that is associated with these products. This is also different between the US and the UK.

In the former case the coverage seems to be a simple predefined percentage of the value of the debt and in the latter case, the loss coverage is shared between lender and insurer.

The Insurers liability is capped to a predefined amount. The figure used is a function of the difference between the LTV ratio of the loan and an arbitrarily defined ‘normal’ LTV ratio (taken as 75%) and so clearly the values of these products do not coincide.

The type and level of the negative incentives to early termination and the presence of insurance coverage may exert a significant role in determining equilibrium contract rates. Consequently even for contracts that agree in every respect except these two, the contract rates would differ and the same would happen with the values of all the underlying assets.

To conclude, the single point that is emphasised here is the inadequacy of the models designed to value US Mortgages when adapted to price British mortgage products and vice-versa.

Under the framework considered in this research, the contracts are in equilibrium; this means that they represent a fair deal. Thus, if in a particular case of a certain contract, the borrower pays more than what he would have paid under a different agreement, this is not taken to imply that the latter economically dominates the former. It is more that the borrower receives greater value in the other components of the contract – the options to terminate the loan. Otherwise it would not be possible for both contracts to represent fair deals.

6.2 Benchmark Parameters

In order to present and discuss the numerical results provided by the model, a basic set of economic parameters was taken into account. The choice made being informed mainly on the basis of the standard assumptions in the literature and on recent mortgage market conditions.

The following table presents these values. Unless otherwise specifically mentioned, this set of economic parameters should be taken to apply.

Table 2: Base values of Parameters for simulations

LIST OF PARAMETERS	**CONTRACT: Repayment Mortgage**		
	25 YR	**5 YR**	**3 YR**
ECONOMIC EVIRONMENT :			
Spot interest rate, r(0) :	6%	6%	6%
Long term average of the interest rate (steady state rate), θ :	7%	7%	7%
Speed of reversion, γ :	25%	25%	25%
House service flow, ϕ :	7.5%	7.5%	7.5%
Correlation coefficient, ρ :	0	0	0
CONTRACT PROVISIONS :			
Maturity, T :	300 months	60 months	36 months
Value of house at origination, H(0) :	£100, 000	£100, 000	£100, 000
Arrangement Fee, ξ :	0.05%	0.05%	0.05%
Early Termination Penalty, π :	1%	1%	1%

As all the parameters are fixed at the moment when each solution is achieved, the only mechanism through which it is possible to search for equilibrium is through a change in the contract rate. Consequently an iterative process is used - based upon Newton’s Method - to determine the equilibrium contract rate. Each parameter exercises an influence on the value of the mortgage and its related assets. Any change in a parameter used to characterise the economic environment leads to a change in the equilibrium contract rate.

As a result, the direct effects resulting from changes of the parameters themselves, the value of the mortgage assets is also influenced by an indirect effect resulting from modification of the contract rate.

6.3 Illustrative Numerical Results

In order to facilitate a general understanding of the way in which numerical results are generated, a description and diagrammatic representation of the evolution of the different options inside the grid that is used to represent the state space is now provided. The four figures based upon the repayment mortgage type as presented here take inspiration from the figures presented in the published work of Kau, Keenan, Muller and Epperson, (1992) and reproduced here.

The motivation for presenting the 2D cross sectional diagrams is to provide a qualitative evaluation of the complex interactions between the embedded options in the mortgage contract. It should be borne in mind that the mortgage valuation surfaces themselves are actually much more complex than portrayed by these diagrams as can be ascertained by their individual discussion, presented in a subsequent section.

There are two main reasons for displaying the 3D figures: in the first place, they show that the solutions evolve across the state space in a smooth way without instabilities. In the second place, they indicate that the shape of the graphs makes sense in economic terms.

Also provided are graphical representations of the component valuations and tabulated sensitivities that are thought to contribute to an economic understanding of complex interrelationships.

The state space, *(H, r)*-space, is compacted into a unit square. Each of the four figures represents a particular moment in lifetime of the mortgage.

Since the solution progresses backwards in time, the sequence of figures proceeds in the same manner. The valuation functions are symbolised as follows:

V = Value of the mortgage for the borrower;

A = Value of the remaining mortgage payments;

P = Value of the American call option to prepay the mortgage at any time to maturity of the mortgage;

D = Value of the European compound put option to default on the mortgage at a monthly payment date;

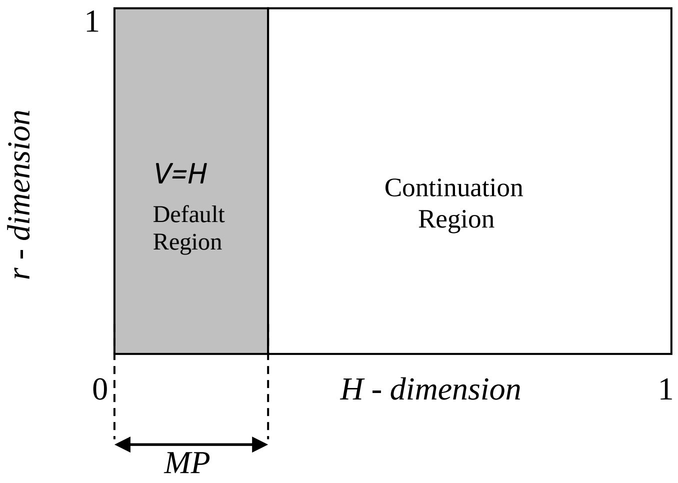
I = Value of mortgage Insurance;

COI = Value of the mortgage Coinsurance.

Termination of the Mortgage Loan: At the final moment in the life of the mortgage it only makes sense for the borrower to make the last monthly payment *MP*, if the value of the house *H* is greater than the amount to be paid, otherwise the borrower will default. If default is not in the best interest of the borrower then the only alternative is to pay.

At termination it is not possible to prepay, as the loan has reached the final moment in its life.

Figure 2: Evolution of the Options to terminate the loan across the state space at different instants in time during the life of the loan: At termination of the loan

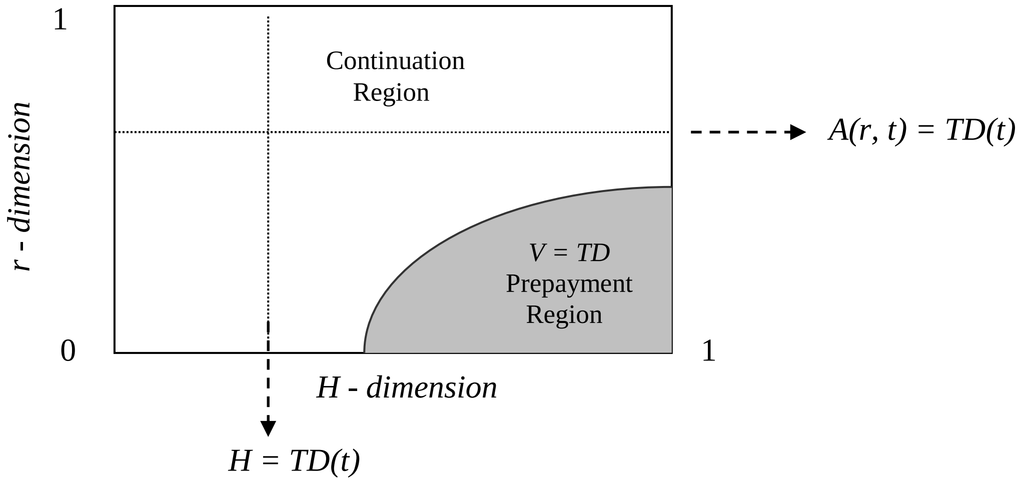


There are thus only two regions in the state space, the default region (shaded) – where the borrower exchanges the value of the last payment for the house – and the continuation region which implies that the last payment is made.

Points in time between the payment dates. In this situation it is the default region that does not exist - default only makes sense at a payment date, since the borrower would lose the service flow for a period of time during which he could enjoy it for free! There are thus only two regions in the state space here too, the prepayment region (shaded) and the continuation region.

Prepayment is a reasonable option if the borrower is facing high house prices, H, and low interest rates, r, at the same time. At high house prices the default option is less valuable and as a result the relative attraction of the prepayment option increases for higher interest rate levels. This is equivalent to saying that the prepayment boundary is positively sloped, given that the borrower is more disposed to prepay at higher interest rates when house prices are high.

Figure 3: Evolution of the options to terminate the loan across the state space at different instants in time during the life of the loan: Between monthly payment dates



The Prepayment region is bounded in the interest rate dimension by the point that corresponds to *A(r, t) = TD(t)* where *TD(t)* represents the total debt at time *t*. If interest rate values go below this level then prepayment would be '*in-the-money'*. However, it is important to note that prepayment does not necessarily take place immediately as the model would suggest. This leads to heterogeneity in borrower prepayment behaviour and the *'burn-out'* effect as evidenced in the literature pertaining to the valuation of mortgage backed securities.

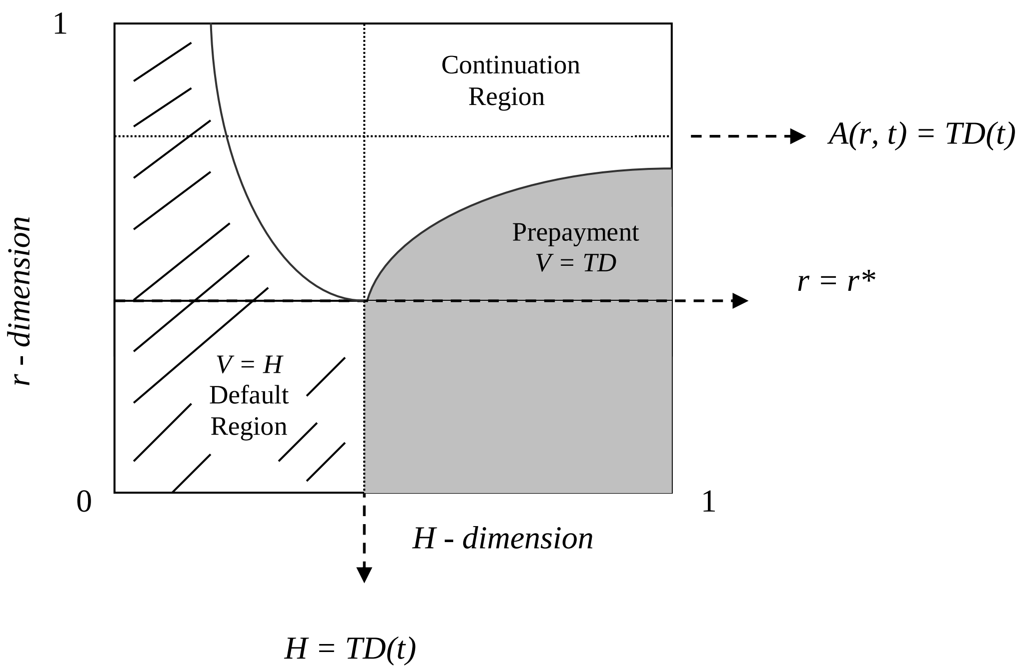
The Prepayment region is bounded in the House price dimension by the point that corresponds to *H = TD(t)*. This represents the borderline in which the relative appeal of default to prepayment is the same. Note however that default never occurs as previously identified.

At intermediate payment dates. At payment dates the default option makes economic sense and so the diagram includes both a prepayment and a default region.

The dividing line between them corresponds to the point where *H =TD(t)*

In financial terms we have that prepayment is as attractive an alternative as default. The highest value of *r* observed along this boundary provides a critical value, since, for rates below this critical value the loan is automatically terminated. Through default if $H < TD$ or through Prepayment if $H > TD$.

Figure 4: Evolution of the options to terminate the loan across the state space at different instants in time during the life of the loan: At monthly payment dates.



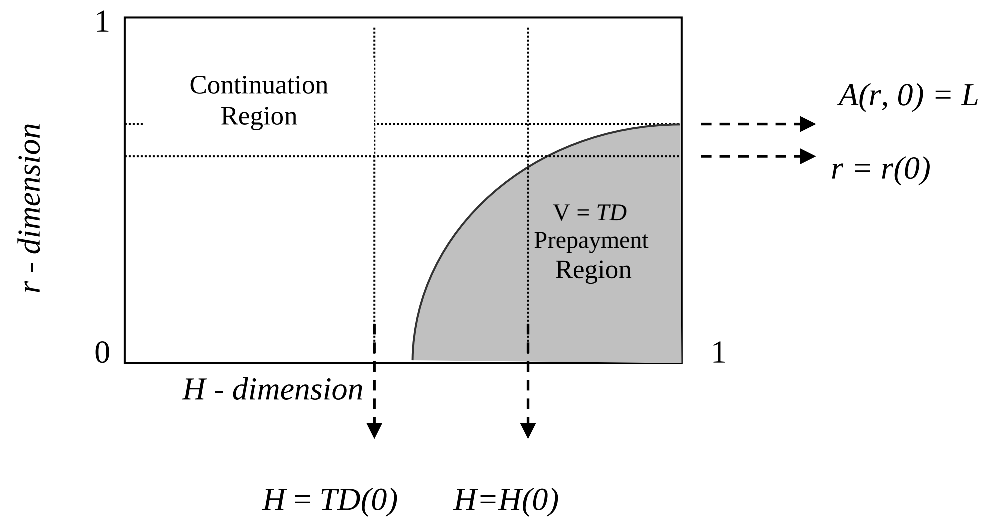
For default to be *'in-the-money'* it is required that the value of the house *H* is less than the present value of the future mortgage payments *A*.

There is a similarity with prepayment, in that exercise of default implies immediate loss of the option to prepay and so default does not happen immediately either, as the model would suggest. Note that high interest rates increase the costs of maintaining the loan and so lead to default being advantageous only at how house prices and as a result the boundary between default and continuation is sloped negatively.

At Loan Origination. There are only two regions in the state space. Prepayment region (shaded) and the Continuation region. There is no default region, as the moment does not correspond to a payment date.

The original combination of state variables corresponds to *(H(0), r(0))*

Figure 5: Evolution of the options to terminate the loan across the state space at different instants in time during the life of the loan: At loan origination



If there is an immediate drop in interest rates, *r*, the borrower may be disposed to prepay the mortgage with the mortgage value *V* declining to the value of the outstanding debt *TD*. Similarly, if there is an immediate rise in house price, *H*, a situation may develop in which prepayment is in the best interest of the borrower as the contract rate is too high.

6.3.1 Overview of the Different Valuation Functions

The valuation functions for each one of the variables under study and a series of figures indicating the behaviour of these functions for different levels of the state variables at the start of the contract are given in this section.

All assets are valued in relation to par value of the house. In each case the figures are based on the entire state space corresponding to a (51, 51) node grid with the centre of the grid located at (26, 26). For the base parameters this corresponds to interest rate 6% and House Price £100 000.

The figures show that the solutions evolve across the state space in a smooth way without instability and indicate that their shapes make sense in economic terms. As is to be expected the main characteristics and the general shapes of the valuation functions do not change substantially between types of contracts. The differences are in the detail and will be discussed in what follows:

Figures have been produced for three contracts:

(i) A 3 year Repayment Mortgage;

(ii) A 5 year Repayment Mortgage and

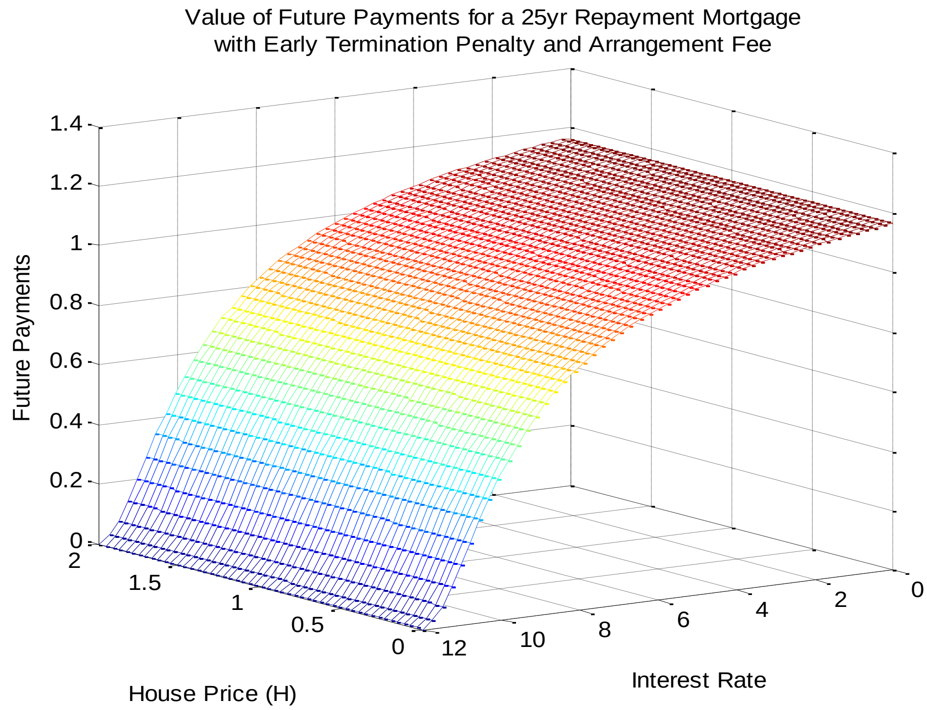
(iii) A 25 year Repayment Mortgage

But only those for the 25 year Repayment Mortgage contract are used in this report for the purposes of illustration as the same pattern is demonstrated across all these contracts.

Future Payments A: The variable *A* depends only upon the interest rate, *r* and therefore the value of the function is constant along the house price dimension *H*.

Figure.6, as to be expected demonstrates that there is an inverse relationship with *r*.

Figure 6: Value of the future payments



The different equilibrium contract rates associated with each mortgage are the cause of the divergence in the value of the mortgage payments that generate small divergence in the value of the function across contracts.

Value of the Mortgage V: This is a rather complex function of the values of *A*, *P*, and *D*. Figure. 7. Illustrates the surface produced.

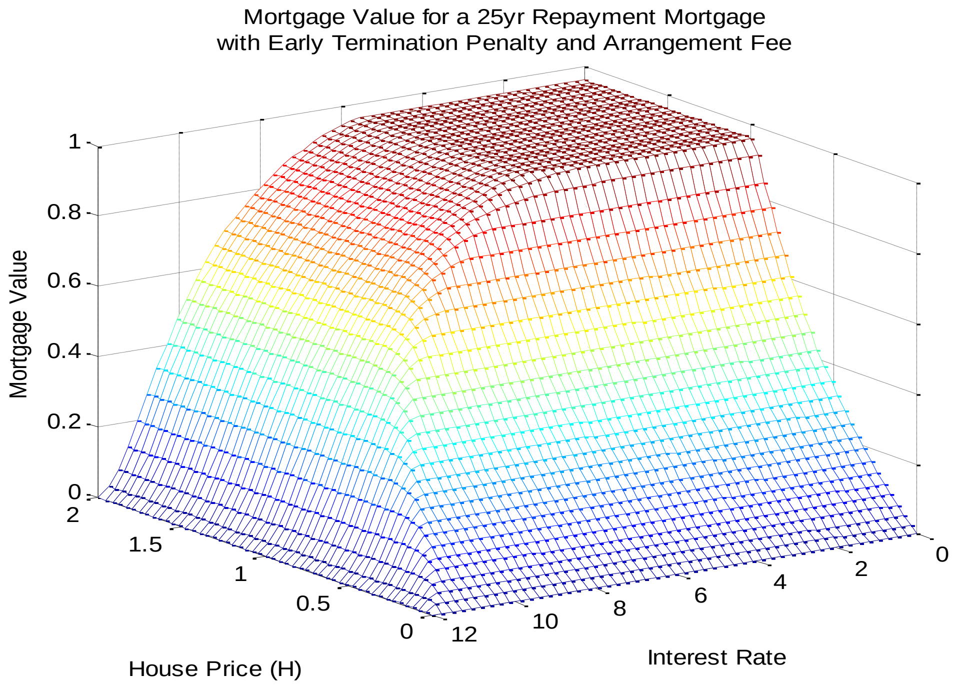
Low levels of House Price, *H* tend to increase the value of the Default Option, *D*, held by the borrower and as a result tend to reduce the value of the mortgage contract, *V*.

The evolution of interest rates impacts more directly *A* and *P*. The effect of this is to produce opposing effects on the value of *V*.

In the former case the relationship is direct and in the latter case there is an inverse relationship to the value of the mortgage. Clearly *P* cannot be greater than *A* and is normally smaller. The primary determinant of the effect of the evolution of interest rates on the mortgage value is provided through *A*

Where Low interest rates and High house prices intersect borrowers prefer prepayment. The prepayment region constitutes a plane levelled at the original value of the loan for contracts with no penalties. With penalties attaching the surface is again a plane but with a level that is now slightly higher.

Figure 7: Value of the mortgage

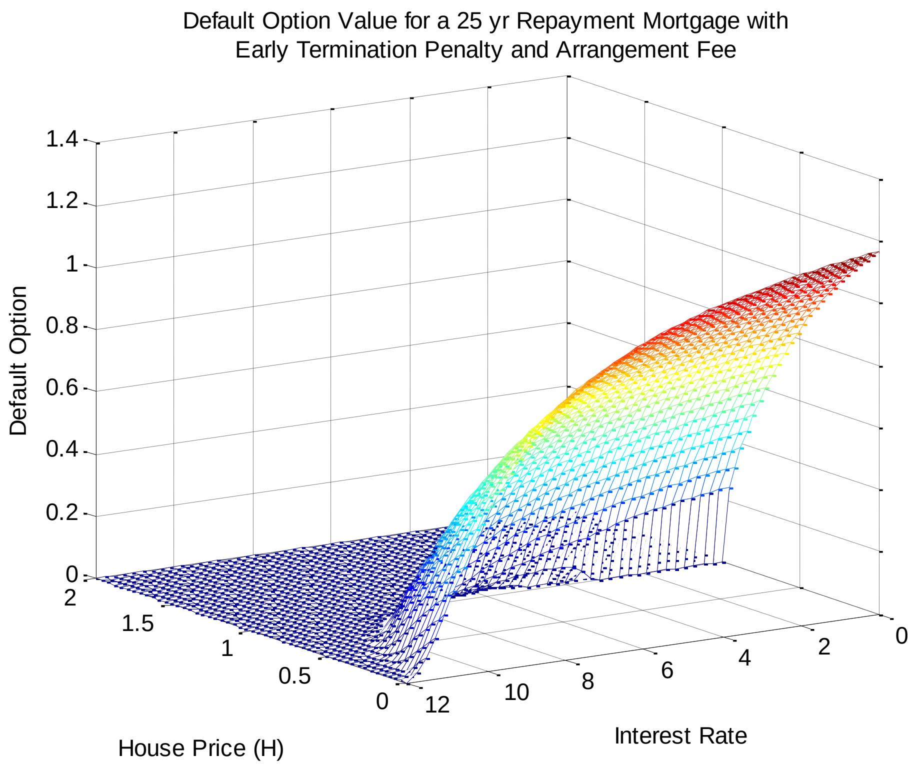


The Default Option D: The primary determinant of the value of this option is the level of *H* and its relation to the value of the mortgage *V*. This is through the relation $V = A - D - P$. Low levels of *H* tend to generate default.

As the increase in the level of *r* leads to decrease in the value of *A* and *V* then the value of the default option, *D*, tends to be inversely related to *r*.

Figure.8. Demonstrates the shape of the default surface.

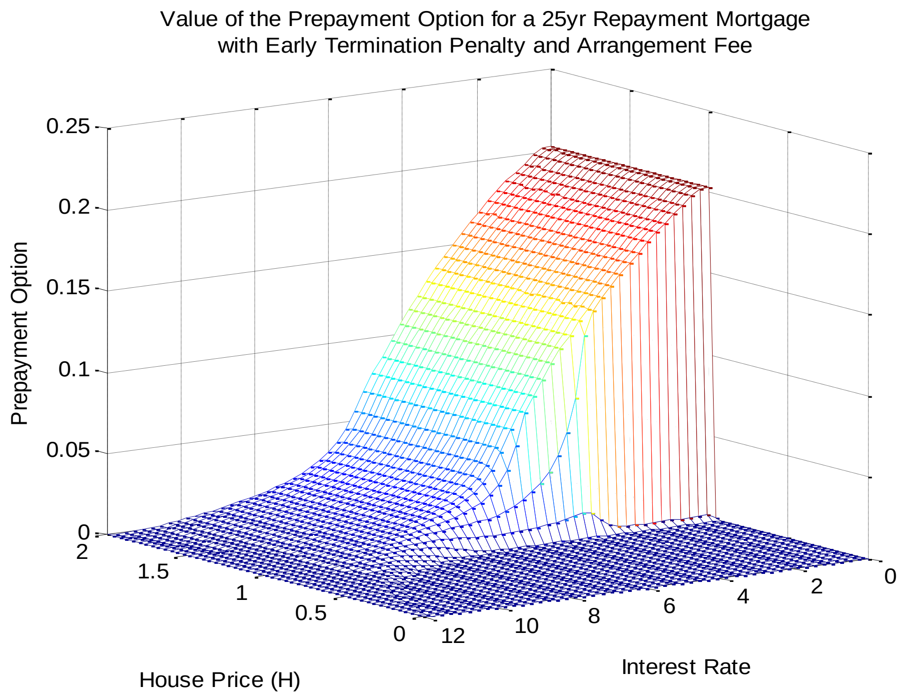
Figure 8: Value of the default option



The Prepayment Option P: The key driver of the value of this option is the interest rate, *r*. The prepayment option value *P*, only takes on high values for low values of interest rate *r* that coincide with high values of house price *H*. This happens because low house prices generate default and a defaulted mortgage cannot be prepaid.

Figure.9. Illustrates the Prepayment option value surface.

Figure 9: Value of the prepayment option



Insurance Products, I and COI: Both of these products values are directly related to the evolution of the default option. When interest rate levels are high the value of the contract *V* is dramatically reduced and house prices need to fall substantially for the borrower to default and the insurance policy to be exercised.

Figure.10 and Figure.11 provide an illustration for the mortgage contract.

The value of the insurance, *I*, is capped so for low house prices the function quickly reaches a maximum with the coinsurance reaching high values only after the cap level is exceeded.

Figure 10: Value of the insurance coverage

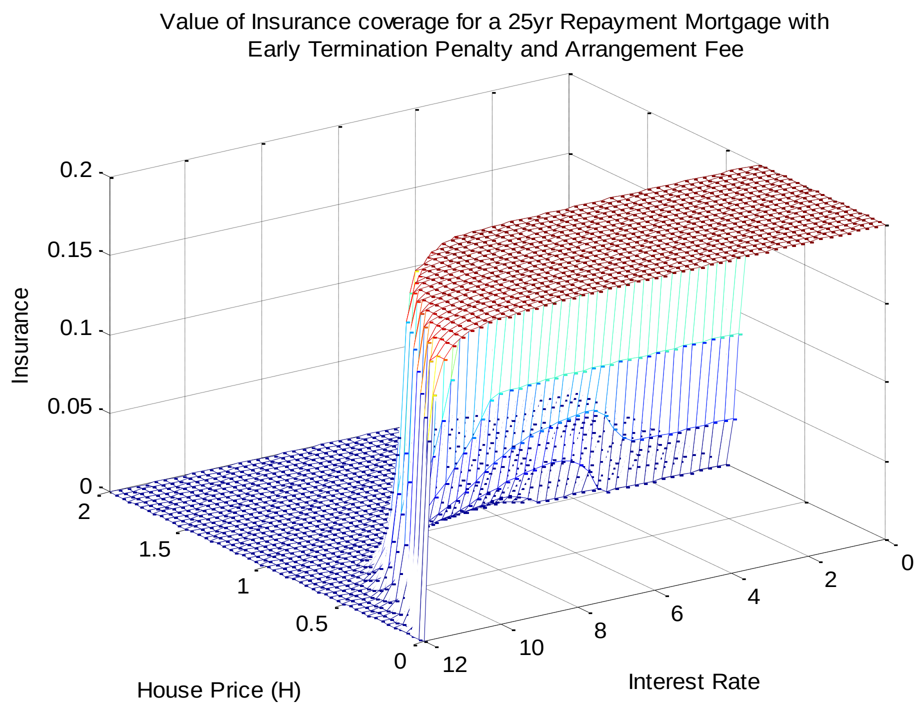
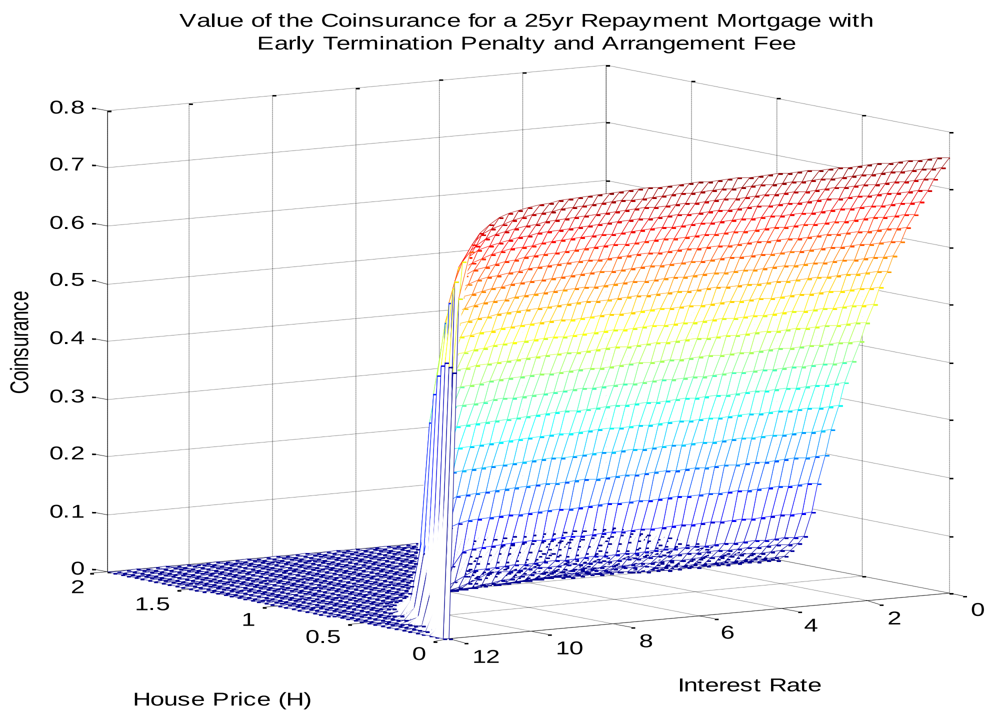


Figure 11: Value of the coinsurance coverage



Equilibrium Contract rates at the Origination of the Mortgages.

This section gives the framework in which the equilibrium contract rates are determined. Mortgage contract rates and contract provisions vary widely over time. The economic environment changes constantly and specifications are subject to frequent revision to accommodate these changes (and for marketing reasons).

In the equilibrium framework adopted here, a mortgage contract can only be acceptable if it represents a fair deal for borrower and lender. Neither should be able to make any *'a-priori'* profit - the no arbitrage condition.

In the model the values of the state variables *H(0)* and *r(0)* are known at origination of the contract so the identification of an equilibrium contract rate is an iterative process. With these values providing the initial starting point along with the contract provisions, a search is performed until the equilibrium contract rate is found. At this rate, the condition of no-arbitrage is met for each party involved. This search routine is implemented by a technique based upon Newton's Iteration method. The search is stopped when the margin of error is within a £10 margin on a £100 000 House (0.0001 of the house value).

Mortgage clauses influence directly the possibility of early termination and the value of the mortgage to the lender can be identified in the arrangement fee ξ, the early termination penalty π and the mortgage Indemnity Insurance, *I*.

The Arrangement fee and Insurance component effects are completely different.

The inclusion of an arrangement fee simply leads to a linear increase in the value of the lender position in the contract, whereas the Insurance component influence is not well defined being quite non-linear for different levels of the contract rate.

At origination, the critical set of values of π, ξ and c that allow

$V(c, \pi) - (1-\xi)L + I(c, \pi) = 0$ correspond to an equilibrium contract.

Most Basic Contract. This corresponds to the situation of a repayment mortgage without Mortgage Indemnity Insurance, Arrangement fees and Early Termination Penalty.

Since these provisions do not exist the equilibrium condition becomes $V(c) - L = 0$**.** In order for this contract to be viable, it is required that the value of the mortgage to the borrower *V*, be equal to the amount lent, *L*.

For this to happen *(H(0), r(0))* sits on the prepayment boundary and immediate prepayment would be a possible optimal strategy for the borrower. Kau et al., (1995).

Table.3 Column *V-L*, and the corresponding line in Figure. 12, Illustrate this situation. No equilibrium exists, even though possible contract rates can be found that are fair deals but these rates correspond to situations of immediate mortgage prepayment.

Table 3: Effects of changes in Contract Rate on the Value of the Mortgage and Mortgage related Assets: Repayment Mortgage without an Arrangement Fee and Early Termination Penalty

(£)

	Future Payments			**Prepayment**	**Insurance**	**Coinsurance**	**Value of the Mortgage to the lender**			
CRATE	**A**	**Mortgage V**	**Default D**	**P**	**I**	**COI**	**V - L**	**V - L + I**	**V - (1 – ξ)L**	**V - (1 – ξ)L + I**
4.0	72516	72087	221	208	486	127	-22913	-22427	-22913	-22427
5.0	80312	79470	609	234	967	242	-15530	-14563	-15530	-14563
6.0	88516	86580	1522	414	1566	392	-8420	-6854	-8420	-6854
7.0	97099	92114	3023	1962	1711	428	-2886	-1175	-2886	-1175
8.0	106034	94790	3272	7974	130	214	-210	-80	-210	-80
9.0	115291	94999	12	20291	0	0	-1	-1	-1	-1
10.0	124840	94999	0	29840	0	0	-1	-1	-1	-1
11.0	134650	95000	0	39653	0	0	0	0	0	0
12.0	144694	95000	0	49698	0	0	0	0	0	0

Table 4: Effects of changes in Contract Rate on the Value of the Mortgage and Mortgage related Assets: Repayment Mortgage without an Arrangement Fee and with an Early Termination Penalty

(£)

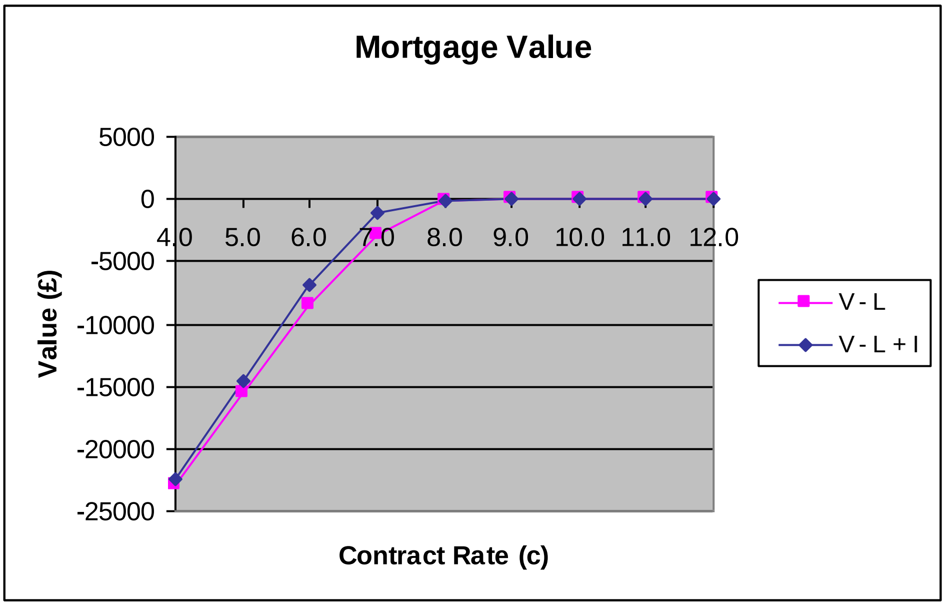
	Future Payments			**Prepayment**	**Insurance**	**Coinsurance**	**Value of the Mortgage to the lender**			
CRATE	**A**	**Mortgage V**	**Default D**	**P**	**I**	**COI**	**V - L**	**V - L + I**	**V - (1 - ξ)L**	**V - (1 - ξ)L + I**
4.0	72516	72097	221	207	486	127	-22903	-22417	-22903	-22417
5.0	80312	79478	609	225	969	243	-15522	-14553	-15522	-14553
6.0	88516	86649	1533	334	1576	394	-8351	-6775	-8351	-6775
7.0	97099	92397	3198	1505	1781	445	-2603	-822	-2603	-822
8.0	106034	95424	4517	6059	990	247	424	1414	424	1414
9.0	115291	95949	24	19341	3	1	949	952	949	952
10.0	124841	95948	0	28892	0	0	948	948	948	948
11.0	134651	95949	0	38701	0	0	949	949	949	949
12.0	144694	95949	0	48748	0	0	949	949	949	949

Table 5: Effects of changes in Contract Rate on the Value of the Mortgage and Mortgage related Assets: Repayment Mortgage with an Arrangement Fee and Early Termination Penalty

(£)

	Future Payments			**Prepayment**	**Insurance**	**Coinsurance**	**Value of the Mortgage to the lender**			
CRATE	**A**	**Mortgage V**	**Default D**	**P**	**I**	**COI**	**V - L**	**V - L + I**	**V - (1 - ξ)L**	**V - (1 - ξ)L + I**
4.0	72516	72087	221	207	498	131	-22913	-22415	-22438	-21940
5.0	80312	79482	609	222	999	251	-15518	-14519	-15043	-14044
6.0	88516	86688	1538	291	1648	412	-8312	-6664	-7837	-6189
7.0	97099	92541	3269	1291	1917	479	-2459	-542	-1984	-67
8.0	106034	95701	5007	5328	1154	288	701	1855	1176	2330
9.0	115291	96424	2248	16625	306	76	1424	1730	1899	2205
10.0	124841	96423	11	28416	1	0	1423	1424	1898	1899
11.0	134651	96422	0	38217	0	0	1422	1422	1897	1897
12.0	144694	96423	0	48271	0	0	1423	1423	1898	1898

Figure 12: Mortgage value: 25yr repayment mortgage without an arrangement fee and early termination penalty

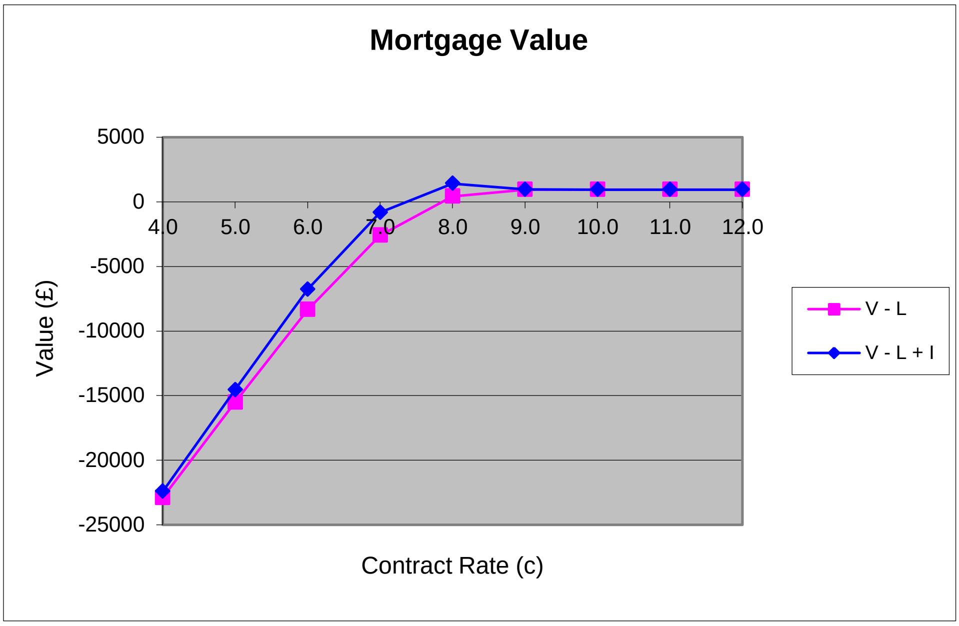


Inclusion of Prepayment Penalties, π. By inclusion of a prepayment penalty the contract equilibrium condition is given by $V(c, \pi) - L = 0$

This situation is revealed in Table. 4. Column *V - L* and the corresponding line in figure. 13. As can be seen, instead of facing an artificial situation of continuous equilibrium, there is only one value of contract rate, *c*, that generates an equilibrium contract.

Introducing the prepayment penalty translates the line upward as the borrower now faces an additional cost. There is now a unique combination of *c* and π that produces equilibrium.

Figure 13: Mortgage value: 25yr repayment mortgage without an arrangement fee and with an early termination penalty



Note, that if the LTV Ratio were 100% then equilibrium combinations would again be impossible. In this case $L = H$ and so the equilibrium condition would become: $V(c, \pi) - L = 0$

By defaulting the borrower loses the service flow of the house and so there can be no contract rate, c, that makes this condition hold.

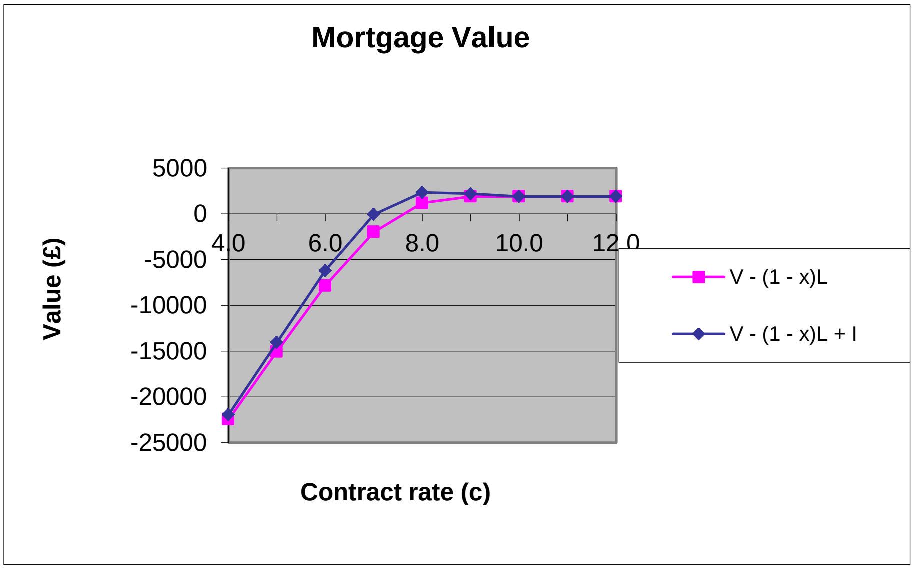
Inclusion of Arrangement Fees. Including arrangement fees adjusts the equilibrium conditions:

$V(c) - (1-\xi)L = 0$ for contracts without an early termination penalty and

$V(c, \pi) - (1-\xi)L = 0$ for contracts with an early termination penalty.

In these cases the curves, represented in Figures. 12. and 13, shift upwards by the amount of the arrangement fee ξL as indicated in Table. 5. Column V-$(1$-$\xi)L$ and corresponding line in Figure. 14.

Figure 14: Mortgage value: 25yr repayment mortgage with an arrangement fee and early termination penalty



As can be seen single equilibrium combinations are reached.

Inclusion of Mortgage Indemnity Guarantees. The effects induced by the inclusion of Mortgage Insurance are different.

For a repayment mortgage without any prepayment penalty or arrangement fee: Initial increases in the contract rate tend to be translated into increases in the value of the Insurance. This is due to the value of *A* getting larger and subsequently the possibility of default increasing.

Similar changes at higher levels of contract rate tend to reduce the value of *I* as they increase the likelihood of early termination by prepayment.

Therefore the situation of continuous equilibrium is reached for high enough levels of contract rate. Under these conditions Insurance has no value as prepayment constitutes the optimum action of the borrower. This situation is revealed in Table. 3. Column *(V-L+I)* and the corresponding line in Figure. 12.

If we next consider mortgage contracts that include early termination penalties and Insurance but do not include arrangement fees as illustrated in Table. 4. Column *(V-L+I)* and the corresponding line in Figure. 13. The inclusion of the penalty masks the effects of the evolution of the contract rate in the value of the insurance and so equilibrium combinations exist

Full Mortgage Contract. This is the case where all the common contractual features are considered and is the case considered in Table. 5. Column $V- (1-\xi)L + I$ and the corresponding line in Figure. 14.

This always results in a situation providing a unique equilibrium combination. The net benefit of these contractual features is to the lender and so combinations are reached at slightly lower levels of the contract rate.

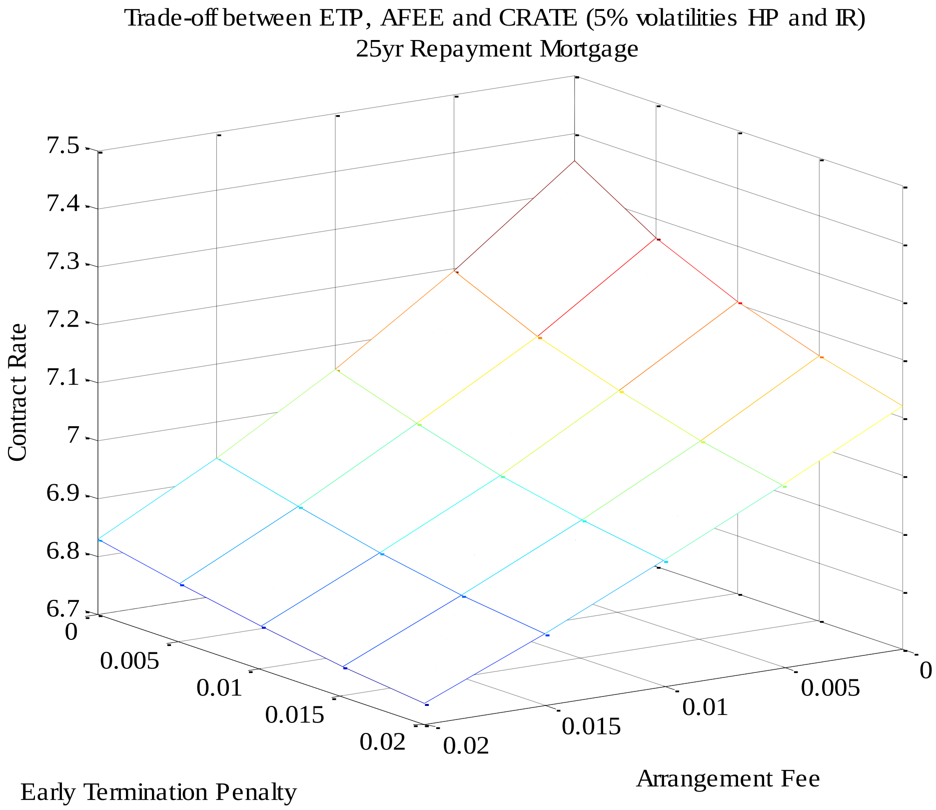
Considering a full mortgage contract with an arrangement fee, ξ, as the only negative incentive to early repayment, the equilibrium contract rate will be a function of the level of this feature so that a curve will exist giving all equilibrium combinations. Considering an early termination penalty, π, adds an extra dimension so that in considering both these negative incentives a surface is created in the (c, ξ, π) space.

Table.6, and Figure.15, demonstrate this situation and provide for an intuitive idea of the trade-offs that exist between the variables for a 25yr repayment mortgage.

Table 6: Trade-off between Early Termination Penalty, Arrangement Fee and Contract Rate for a 25yr Repayment Mortgage

Early Termination Penalty ETP	Arrangement Fee AFEE				
	0.000	**0.005**	**0.010**	**0.015**	**0.020**
0.000	7.35	7.20	7.06	6.94	6.83
0.005	7.27	7.13	7.01	6.90	6.80
0.010	7.20	7.08	6.97	6.87	6.77
0.015	7.16	7.05	6.94	6.84	6.75
0.020	7.12	7.02	6.92	6.82	6.74

Figure 15: Trade-off between early termination penalty, arrangement fee and contract rate



6.3.2 Sensitivity Analyses.

This section details brief analysis of the effects induced by changes in the parameters used to characterise the economic environment along with an analysis of the consequences of the effects induced by changes in the parameters of the mortgage contract.

Effects induced by changes in the Economic Environment

- *Volatility of state variables*

Here we are going to examine the effects of the risk created by changes in the state variables. For a previously determined equilibrium contract rate, the risk parameters are changed so allowing an assessment of the partial effects induced by changes in volatilities of the state variables.

- *Interest Rate Volatility*

Interest rate volatility affects the values of all the mortgage-related assets. Table.7. and the corresponding Figure. 16. Illustrates the simulation results.

Table 7: Effects induced by changes in Interest Rate Volatility on the Value of the Mortgage and Mortgage related Assets

25yr Repayment Mortgage (£)

Interest Rate Volatility	**Future Payments**	**Mortgage Value**	**Default**	**Prepayment**	**Insurance**	**Coinsurance**
Sigma	**A**	**V**	**D**	**P**	**I**	**COI**
0.050	97811	92746	3347	1719	1775	444
0.075	101404	92541	4595	4020	1737	434
0.100	105934	91352	5913	7165	1668	417
0.125	111657	90171	7371	11244	1482	370
0.150	118452	89041	8937	16229	1229	307

The value of the Future Mortgage Payments increases with increase in interest rate volatility. The reason for this result is that two changes of the same absolute size, but of opposite sign, in the value of the discount rate used to determine the present value of a future cash flow, results in changes of different size in the present value itself.

That is, the gains observed with a fall in discount rate exceed the losses observed with a rise of the same magnitude. The result is an increase in the expected value of future payments; this effect is not dependent in any way upon the type of contract. The differences in the values of A observed in the table are thus determined by the different contract rates that are found for each row.

All the other mortgage related assets represent future cash flows that are either directly or indirectly affected by A, through the relationship $V = A - D - P$

Although this effect pertains to them all, its influence may not always be significant in determining the relationship between the mortgage asset and the interest rate volatility.

For the case of the Prepayment Option, P an increase in volatility of interest rates will tend to increase the possibility that the contract reaches the prepayment region in preference to either continuation or default.

Prepayment is also a function of A and so the combined result of this is that the Prepayment Option values tend to move in direct relationship with changes in volatility of interest rates.

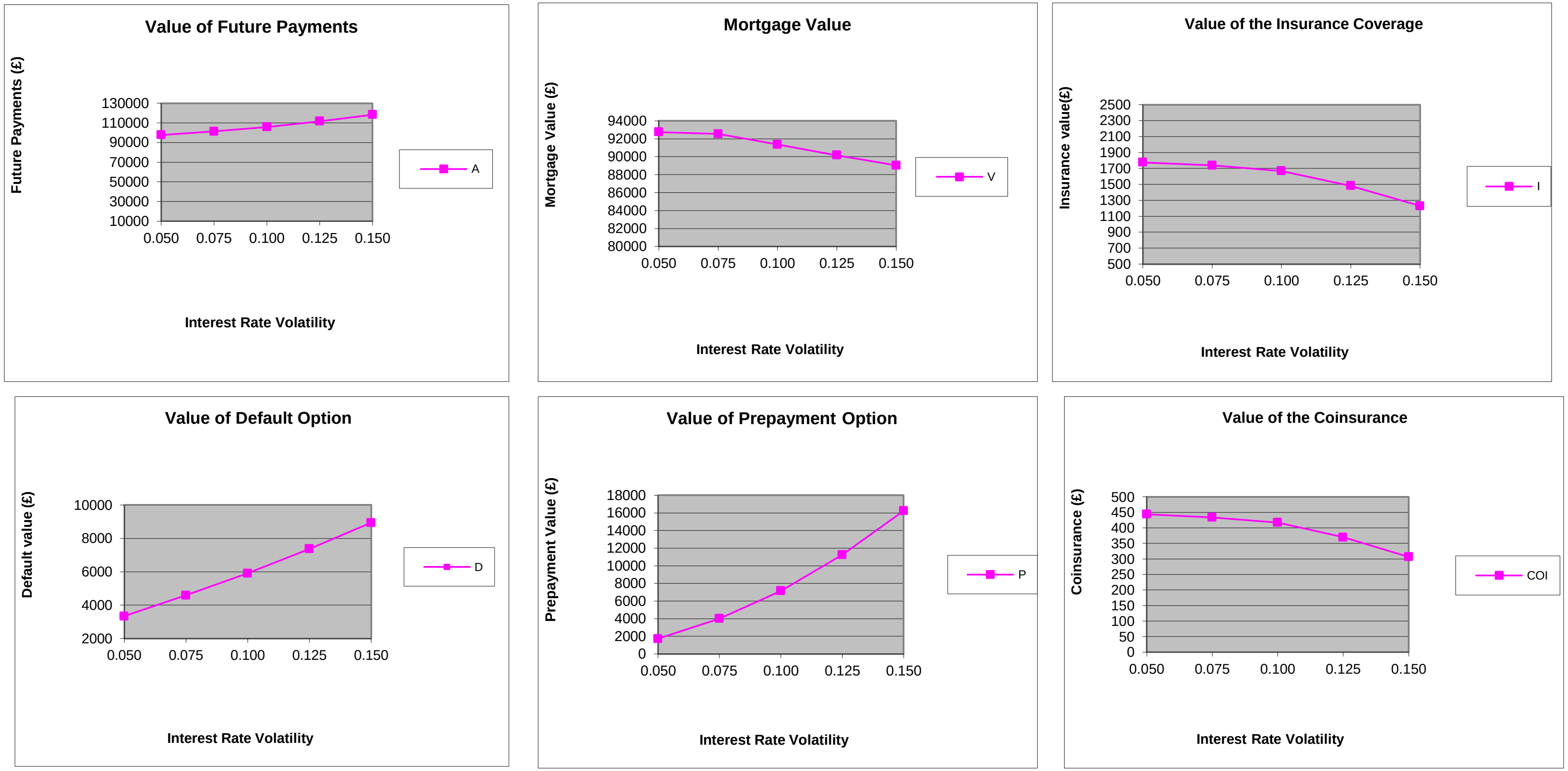
In the case of the Default Option, *D*, the impact of changes in interest rate changes the relative size of the different regions within the grid and as such this impact favours a negative relation between *D* and the volatility in interest rates. The default option is also a function of *A*. The default option dominates in any relative changes of *P* in relation *D* with the result that *D* also tends to move in direct relationship with changes in volatility of interest rates.

The Mortgage related insurance assets *COI* and *I*, both tend to move in an inverse relation with changes of interest rate volatility. This is due to increases in interest rate volatility tending to translate into greater prepayment levels and so there is a reduction in default level.

Since *COI* and *I* do not benefit from any change in *A*, the loss is determined by the outstanding balance so these assets values hold an inverse relationship with change in interest rate volatility.

As a result the value of the lender's position evolves in an inverse relationship with interest rate volatility and the value of the *COI* and *I*, present negative relationships with interest rate volatility.

Figure 16: Effects of changes in Interest Rate Volatility on the Value of the Mortgage and Mortgage related Assets



- *House Price Volatility.*

House Price volatility affects the mortgage-related assets in a significantly different way. Table.8. and the corresponding Figure. 17. illustrate the simulation results.

Table 8: Effects induced by changes in House Price Volatility on the Value of the Mortgage and Mortgage related Assets

25yr Repayment Mortgage (£)

House Price Volatility	**Future Payments**	**Mortgage Value**	**Default**	**Prepayment**	**Insurance**	**Coinsurance**
Niu	**A**	**V**	**D**	**P**	**I**	**COI**
0.050	97099	92541	3269	1291	1917	479
0.075	97099	91243	4752	1105	3147	787
0.100	97099	89652	6503	945	4563	1143
0.125	97099	87901	8356	843	6144	1569
0.150	97099	86059	10283	758	7495	2159

The value of Future Payments, *A*, is constant - it is derived from the interest rate, *r*. This means that *A* is independent of House price, *H*, and house price volatility.

In the case of Default, *D*, there is evidence of a direct relationship with house price volatility. Increase in house price volatility tends to preferentially increase likelihood of default. With the direct result that Prepayment *P* decreases with increases in house price volatility.

Now, the relationship with the value of the Mortgage, *V*, is determined by the effect of House Price volatility on the joint option to terminate, *(D+P)*. Such changes in the relative values of *D* and *P* tend to produce opposing effects on the joint option to terminate. Since the effects of *D* dominate the effects of *P* the result is that the mortgage value, *V*, tends to present a negative relationship with House Price volatility.

The Mortgage related Insurances, *COI* and *I*, display a direct relationship with House price volatility. House Price volatility impacts upon default in a much greater way than prepayment and as the value of these assets are a direct function of the likelihood of default, the evolution of these assets is thus very much akin to the evolution of default.

Figure 17: Effects of changes in House Price Volatility on the Value of the Mortgage and Mortgage related Assets

- *Global Effects.*

An analysis of Table 9 shows the total effects of changes in the volatility parameters within each level of LTV ratio, giving insight into the effects at this level.

The effects caused by interest rate volatility change are straightforward. Increase in interest rate volatility, σ, leads to growth in Future payments, *A*, and so generates increase in Default, *D* and Prepayment *P*.

Insurance *I*, tends to move oppositely to changes in σ, the resulting global effect of the position of the lender is dependent upon Mortgage Value *V*, and *I*. Both of these tend to decrease with increase in σ and so the contract rate needs to increase to obtain equilibrium combinations. Consequently the value of *A* rises to compensate for the decrease in *I* and the increase in the joint option to terminate the loan, (*D*+*P*).

Changes in the house price volatility are more complex. Increases in House Price Volatility, ν, tend to correspond to increases in *D* and decreases in *P*. The joint effect of this is a reduction in *V*.

I tends to increase with ν. The value of *P* tends to exceed the value of *D* unless the LTV ratio is at a high level, which is an effect that is in agreement with Kau, Keenan, Muller and Epperson, 1995 for the American situation.

- *Different Levels of LTV Ratio.*

To discuss the effects at the various levels of the Loan to Value ratio, LTV ratio, Table 9 is again used, looking at the columnar contents in the body of the table. Here for a given house price, a rise in LTV ratio corresponds to an increase in the amount of the loan *L*, lent to the borrower.

Future Payments *A*, therefore tends to grow as a result of the growth in LTV due to the increase in loan amount.

The likelihood of Default, *D*, tends to increase with a rise in the relation between the original House Price and the amount of *L*. This makes it easier for the level of outstanding debt to be greater than *H* and so leads to a positive relation between *D* and LTV.

This effect is mediated to an extent by the evolution of *P*. Increases in LTV tending to raise the prospect of prepayment occurring and so diminish the prospect of default. For low LTV levels *P* tends to have a direct relationship to *A*.

For high LTV levels the likelihood of default and the default option value increases so that prepayment option value decreases. Under these circumstance as to be expected the value of the insurance related assets *I* and *COI* follow the evolution of *D* and so they exhibit a direct effect with LTV.

The overall effect generated by changes in LTV is determined the how *V* evolves relative to I. For low LTV the size of the change tends to be slightly greater for *I* than *V*. Consequently the contract rate tends to reduce slightly with increase in LTV. For Higher LTV the effects change due to the increased likelihood of default and so here the contract rate sees a slight increase.

As the relative magnitudes of these changes are very similar for both *V* and *I* the effect on the contract rate is in matter of fact minimal.

Table 9: The Combined Effects of changes in LTV Ratios and House Price and Interest Rate Volatilities

(25yr Repayment Mortgage with AFEE and ETP) (£)

LTV	Volatility	Equilibrium Contract RATE		Future Payments A		Mortgage Value V		Default D		Prepayment P		Insurance I		Coinsurance COI	
	House Price	ν=0.05	ν=0.10	ν=0.05	ν=0.10	ν=0.05	ν=0.10	ν=0.05	ν=0.10	ν=0.05	ν=0.10	ν=0.05	ν=0.10	ν=0.05	ν=0.10
	Int. Rate														
95%	σ=0.05	7.08	7.14	97811	98289	92746	90122	3347	6848	1719	1320	1775	4404	444	1102
	σ=0.10	7.82	7.78	105936	105574	92859	90354	5908	10148	1770	5074	1667	4171	417	1046
90%	σ=0.05	7.06	7.08	92483	92621	88634	86408	1688	4767	2162	1448	912	3143	228	808
	σ=0.10	7.88	7.75	100929	99744	88913	86771	2678	7009	9339	5966	641	2784	160	723
85%	σ=0.05	7.07	7.10	87454	87649	84179	82644	765	3279	2512	1727	390	1933	99	721
	σ=0.10	7.92	7.79	95676	94559	84354	82946	975	4468	10348	7146	226	1630	57	556
80%	σ=0.05	7.09	7.21	82460	83284	79461	78838	302	2132	2698	2315	136	761	52	872
	σ=0.10	7.95	7.99	90282	90548	79536	79070	288	2390	10458	9089	64	53	23	585

- *Changes in Correlation Coefficient between both state variables*

Table.10. Presents the effect of changes in the correlation coefficient, ρ, on the value of the Mortgage and Mortgage related assets.

Table 10: Effects induced by changes in Correlation Coefficient

25yr Repayment Mortgage (£)

RHO	CRATE	Future Payments A	Default D	Prepayment P	Insurance I	Coinsurance COI
-0.20	10.835	100259	1376	4786	429	107
-0.15	10.839	100289	1464	4744	436	109
-0.10	10.849	100366	1566	4719	445	111
-0.05	10.853	100394	1646	4681	449	112
0.00	10.864	100472	1749	4654	458	115
0.05	10.867	100496	1837	4605	465	116
0.10	10.877	100572	1927	4589	470	118
0.15	10.880	100591	2010	4540	476	119
0.20	10.890	100669	2106	4517	481	120

The tabulated values demonstrate that increase in ρ, tends to be translated into increase in the value of default indicating a direct relationship with the evolution of interest rate r and House Price *H*. When default *D* is high the likelihood of Prepayment *P* is lowered.

Prepayment moves in opposition to changes in ρ. Given this situation the values of the insurance assets *COI* and *I* follow the trend established by the default option.

Since the size of the change in *D* is greater than in *P* the value of the mortgage *V* will decrease. The movements in *I*, (and *COI*), are minimal.

6.3.3 Conclusions

In summary, The numerical results of the simulations presented here confirm the qualitative reasoning behind the four figures presented earlier in section 5.3 from the work of Kau, Keenan, Muller and Epperson, (1992) these numerical results also suggest that in modelling mortgages one should avoid the exclusion of contractual provisions in order to prevent misleading conclusions.

It has been demonstrated that it is worthwhile to model the value of mortgages with as many provisions included as possible in order to obtain values of the associated assets that are capable of providing a reasonable view of the whole mortgage contract. The practice - that is common in parts of the literature over several decades - of excluding one of the embedded options, in order to simplify the modelling process, may lead to potentially equivocal values and consequently should be used with caution.

It is to be noted that even with recent improvements in raw computing power - since the earlier works of Kau, Keenan, Muller and Epperson, (1995) and Azevedo-Pereira et al. (2000) using finite-difference schemes - fully numerical calculations remain ponderously slow. The most recent work of Sharp et al. (2008) using perturbation methodology does however provide speedy and fairly accurate mortgage valuation approximations.

Chapter 7
Modifications of the Standard Model

7.1 Introduction

The preceding chapters have considered in some detail a realistic mortgage valuation model - the standard two-state variable model - (including the potential for early prepayment and the risk of default), based on stochastic house-price and interest-rate processes. This chapter now considers several modifications to the standard model indicated below

(i) Including Mean Reversion in House Prices

(ii) Modelling Prepayment with a Barrier Option

(iii) Incorporating Current Economic Conditions

and takes a fresh look at fundamentals for a modelling framework that can be used to assess the credit risks inherent in a turbulent market environment

(iv) Including the traditional 3C's of credit underwriting

The issues addressed in this chapter will follow a two-state variable model, using UK Fixed Rate Mortgages as the practical example and will seek to

(i) bridge an apparent gap between the theoretic and empirical mortgage literature that appears not to have been previously given attention,

(ii) confirm recent published research activity into prepayment modelling,

(iii) appeal to a mechanism that will allow the effects of the inclusion of macroeconomic variables to be potentially investigated and

(iv) suggest that a framework that links causes and effects of mortgage termination should be considered particularly in light of recent turbulence in the financial markets.

7.2 Including Mean Reversion in House Prices

Until now, all option theory based mortgage valuation models in the literature have assumed that house prices follow a geometric Brownian motion process (Kau, Keenan, Muller and Epperson, 1992) often termed a random walk. Empirical evidence however suggests that house prices are poorly approximated by a geometric Brownian motion and instead are more consistent with a mean reverting process Meese and Wallace, (1997, 1998). Micro-economic theory also provides a strong support for the notion that the price of a good (in this case housing) in a competitive market is mean-reverting, due to supply and demand responses to prices. For instance, an increase in demand increases prices until a supply response is reached whereby prices are driven back down to the production cost plus a normal economic return.

In this section an extension to the existing random walk model is proposed by allowing house prices to follow a geometric mean reverting process. Here the primary goal is to investigate the implications of a mean reverting house price process has on mortgage valuation, by undertaking a comparison of the results to those of the standard model that assumes a geometric Brownian motion process for house prices.

By creating a theoretical model based upon mean reversion, this section proposes that this is both more plausible and useful. The model as presented suggests additional factors that may help forecast mortgage termination that has been previously overlooked in the literature. More specifically the models implications are that certain elements of local house prices matter for mortgage valuation – for instance deviations from the long run expected house values and the rate of mean reversion in the local housing market. This is found to be somewhat consistent with the results of Lekkas, Quigley and Van Order (1993) who found that geographic region was empirically significant.

7.2.1 Background Preliminaries

The reason mean reversion influences the value of a mortgage is that the value of living in the house (often called the service flow) turns out to be mean reverting. That is, the required rate of return on an investment in a house is composed of the service flow and the expected capital gains, which is mean reverting by assumption in the proposed model. The required rate of return on the investment in the house is exogenously determined by the equilibrium rate of return on similar investments. Since the required rate of return is externally determined independently of the house price process, the service flow is by definition mean reverting (This insight is described in Dixit and Pindyck (1994) in the context of Real Options). Thus the parameters of the mean reverting process will influence mortgage value through the service flow.

The default option is more valuable under the mean reversion model than the standard random walk model if house prices are substantially below the long term trend. Intuition for why the option values differ across the two models comes from thinking about what happens if the house value declines below the value of the outstanding balance. If house prices follow a random walk, then the house value is just as likely to fall as to rise. In the mean reversion model the option value depends upon whether or not the current house value is above or below the long term trend in house value.

If the house value is below its long term expected value, then the value of the house is expected to rise in the mean reversion model, thus making the default option less valuable. Likewise if the house value is above its long term expected value, then the value of the house is expected to fall in the mean reversion model, thus making the default option more valuable.

Since the primary task is to provide a comparison of the implications of mean reversion in house prices to the implications of the standard geometric Brownian motion model attention is restricted to financially motivated prepayment and default only and complete financial markets and rational borrower behaviour are assumed. Thus Default will only occur at the payment date and Prepayment can occur at any time, for low enough interest rates and high enough house values. For good measure simplicity transactions costs and mortgage insurance are included in the results tabulation and only fixed rate fully amortising contracts are considered Simulation results being illustrated by the 25 yr Repayment Mortgage contract for reasons previously identified.

7.2.2 The Mean Reversion Model

Using the notation previously specified the mean reverting model utilises the following stochastic processes:

Interest rate process:

$$dr = \gamma(\theta - r)dt + \sigma_r \sqrt{r} dz_r$$

House price process:

$$\frac{dH}{H} = \eta(\bar{H} - H)dt + \sigma_H dz_H$$

Correlation:

$$dz_r dz_H = \rho dt$$

where the standard Wiener processes are such that E[dz]=0 and E[dz^2]=dt.

and the variables not previously defined in Chapter 3 are:

η = the speed of mean reversion in house prices and

$\bar{H}$ = the long run mean of house prices – assumed to be constant for simplicity, but could equally be generalised to be a function of local regional economic variables.

One difference from the geometric Brownian motion model is that the return on housing reverts at speed η to its long run average, $\bar{H}$. The other difference is that the service flow is a mean reverting function instead of a constant. This is because the required return on the investment in the house μ is composed of the service flow, $\emptyset$, and the expected capital gains

$$\mu = \emptyset + \eta(\bar{H} - H)$$

Rearranging this gives the only difference between the geometric Brownian motion and mean reversion PDE's

$$\emptyset = \mu - \eta(\overline{H} - H) \qquad \text{Eqn (6.1)}$$

According to the Capital Asset Pricing Model, the required return on investment in the house μ should only reflect the assets non-diversifiable risk.

The key difference is that the service flow is now a function of the house price, whereas in the geometric Brownian motion model it is a constant.

This unfortunately replaces the unknown parameter $\emptyset$ with the unknown parameters: $\overline{H}$, η and μ. The parameters can be empirically estimated, but in one of the following sections simulation results for a wide range of values are given.

7.2.3 Model Dynamics

In this section the PDE that describes how the option values evolve over time is presented. Its derivation follows exactly the same arguments as the standard model derived in an Appendix but with the result that the service flow $\emptyset$ is replaced by

$$\mu - \eta(\overline{H} - H)$$

Hence the mean reverting house price model has the following fundamental PDE:

$$\frac{\partial V}{\partial t} + \frac{1}{2}H^2\sigma_H^2\frac{\partial^2 V}{\partial H^2} + \frac{1}{2}r\sigma_r^2\frac{\partial^2 V}{\partial r^2} + \gamma(\theta - r)\frac{\partial V}{\partial r} + (r - \mu + \eta(\overline{H} - H)H\frac{\partial V}{\partial H} + \rho H\sigma_H\sigma_r\sqrt{r}\frac{\partial^2 V}{\partial r\partial H} = rV$$

... ... Eqn (6.2)

7.2.4 Mean Reversion Model Simulation Results and Discussion

This section compares the optimal default and prepayment boundaries from the standard traditional model, labelled 'Tradit' of Chapter 3 and the Mean Reversion model adaptation, labelled 'Mrev' presented here and investigates how sensitive the models are to the parameter values.

To use the models estimates of the parameters of the underlying mortgage value processes are required.

Table 11 below presents the base values of these parameters utilised in the numerical solutions. These values previously identified are assumed to hold here unless otherwise stated.

Table 11: Base Parameters for mean reversion model

LIST OF PARAMETERS	**CONTRACT: Repayment Mortgage**		
	25 YR	**5 YR**	**3 YR**
ECONOMIC EVIRONMENT :			
Spot interest rate, r(0) :	6%	6%	6%
Long term average of the interest rate (steady state rate), θ :	7%	7%	7%
Speed of reversion, γ :	25%	25%	25%
Volatility of interest rate σ_r:	5%	5%	5%
House service flow, ϕ :	7.5%	7.5%	7.5%
Volatility of House Prices, σ_H:	5%	5%	5%
Mean of house Prices H-bar:	£150,000	£150,000	£150,000
Speed of reversion of house Price η:	0.0003	0.0003	0.0003
Correlation coefficient, ρ :	0	0	0
CONTRACT PROVISIONS :			
Maturity, T :	300 months	60 months	36 months
Value of house at origination, H(0) :	£100, 000	£100, 000	£100, 000
Arrangement Fee, ξ :	0.05%	0.05%	0.05%
Early Termination Penalty, π :	1%	1%	1%

Based upon the parameter values above Table 12 shows that the default and prepayment option values under the mean reversion model, are lower in magnitude for the given values of the LTV ratio used in the simulation.

Table 12: Effect on Mortgage Assets of changes in Loan to Value Ratio LTV

(25yr Repayment Mortgage with AFEE and ETP) (£)

Model	LTV Ratio	Contract RATE	Future Payments A	Mortgage Value V
Mrev	**0.80**	7.112	82593	79598
Tradit	**0.80**	7.112	82593	79596
Mrev	**0.85**	7.112	87755	84571
Tradit	**0.85**	7.111	87751	84561
Mrev	**0.90**	7.112	92917	89523
Tradit	**0.90**	7.110	92901	89480
Mrev	**0.95**	7.112	98079	94333
Tradit	**0.95**	7.114	98102	94236

Model	Default D	Prepayment P	Insurance I	Co-Insurance COI
Mrev	2	2993	0	0
Tradit	6	2991	2	1
Mrev	9	3175	3	1
Tradit	29	3161	13	3
Mrev	61	3333	29	7
Tradit	154	3267	69	17
Mrev	522	3224	191	48
Tradit	795	3071	282	70

This stems from the fact that in the mean reverting model, the likelihood of house prices falling so far as to wipe out all of the equity in the house is much less than in the traditional model of Chapter 3, provided of course that the house value is significantly larger than the outstanding balance on the mortgage loan.

Table 13: Effect on Mortgage Assets of changes in House Price Volatility

(25yr Repayment Mortgage with AFEE and ETP) (£)

σ_H	Default		Prepayment	
	Tradit Model	**Mean Reverting**	**Tradit Model**	**Mean Reverting**
0.050	3269	2305	1290	3567
0.075	4752	2741	1105	3242
0.100	6503	3205	945	3115
0.125	8356	4622	843	2674
0.150	10283	7366	758	1753
σ_H	**Insurance**		**Co-insurance**	
	Tradit Model	**Mean Reverting**	**Tradit Model**	**Mean Reverting**
0.050	1917	1258	479	320
0.075	3107	2098	787	527
0.100	4533	3042	1143	768
0.125	6144	4096	1569	1037
0.150	7475	5351	2159	1438

Table 13 shows the value of the default and prepayment options in the mean reversion model for the various values of the volatility in house price.

The effects induced by the house price volatility are more or less straight forward. Increase in volatility leads to simultaneous increase in default and decrease in prepayment. The insurance components tend to move in the same direction as to the movements in volatility.

The positive effect on value, generated by increases in house price volatility, upon default tends to be of higher magnitude than those produced upon prepayment and insurance.

Table 14: Effect on Mortgage Assets of changes in Speed of House Price Reversion (25yr Repayment Mortgage with AFEE and ETP) (£)

Speed of House Price Reversion η	**Default**	**Prepayment**	**Insurance**	**Coinsurance**
0	1719	3347	1775	444
0.0001	1697	3304	1219	299
0.0002	1676	3263	838	202
0.0003	1656	3224	576	136
0.0004	1637	3187	395	92
0.0005	1619	3152	272	63
0.0006	1602	3118	187	42
0.0007	1585	3087	128	28
0.0008	1566	3049	88	19

Table 14 shows the value of the default and prepayment options in the mean reversion model for the various values of the rate of house price reversion. If the speed of the reversion in house prices differs across regions, then it is the local rate of reversion that matters.

Table 15: Effect on Mortgage Assets of changes in required return on the property (25yr Repayment Mortgage with AFEE and ETP) (£)

Mu	**Default**		**Prepayment**	
	Tradit Model	**Mean Reverting**	**Tradit Model**	**Mean Reverting**
0.050	661	455	3176	3270
0.075	795	522	3071	3222
0.100	1003	606	2925	3163
0.125	1810	718	2415	3087

Mu	**Insurance**		**Co-insurance**	
	Tradit Model	**Mean Reverting**	**Tradit Model**	**Mean Reverting**
0.050	234	169	59	42
0.075	282	191	70	48
0.100	365	218	91	55
0.125	804	254	201	64

Table 15 shows the value of the default and prepayment options in the mean reversion model for the various values of the required return on the property.

Giving some indication as to the sensitivity of the mean reversion model compared to the traditional model.

Table 16: Effect on Mortgage Assets of changes in Long Term Mean House Price

(25yr Repayment Mortgage with AFEE and ETP) (£)

HBAR	**Default**		**Prepayment**	
x (£100,000)	**Tradit Model**	**Mean Reverting**	**Tradit Model**	**Mean Reverting**
1.3	1810	713	2415	3080
1.4	1003	604	2925	3154
1.5	795	522	3071	3211
1.6	661	457	3176	3278
1.7	567	403	3235	3317
1.8	493	360	3299	3350
1.9	435	325	3351	3395
2.0	387	295	3399	3417

HBAR	**Insurance**		**Co-Insurance**	
x (£100,000)	**Tradit Model**	**Mean Reverting**	**Tradit Model**	**Mean Reverting**
1.3	804	254	201	64
1.4	365	218	91	55
1.5	282	192	70	48
1.6	234	169	59	42
1.7	205	151	51	38
1.8	180	136	45	34
1.9	161	123	40	31
2.0	144	113	36	28

Table 16 shows the value of the default and prepayment options in the mean reversion model for the various values of the Long Term Mean House Price. According to the numerical solutions, higher levels of house price in relation to the long term mean of house prices tend to be opposite in their effects on the evolution of default and prepayment and also tend to partially compensate each other.

7.2.5 Concluding Remarks

A modified mortgage prepayment and default model has been presented that bridges a divide between theory and empirical valuation that appears to have been overlooked in the literature by challenging the traditional assumption that the service flow is a constant proportion of the value of the house, (see for example, Hendershott and van Order (1987) and Downing (1998)). In this context the service flow is assumed to be a value of living in the house, whilst in the context of options theory it might be modelled akin to a dividend.

The model presented suggests that additional factors relating to the local housing market – in particular deviation from the long run expected house value and the rate of reversion toward the mean house price in the local housing market – may help better forecast mortgage termination and create a more realistic mortgage valuation model.

7.3 Modelling Prepayment with a Barrier Option

The issue addressed in this section is that of ruthless prepayment and the effects of a relaxation in this assumption. Termination within a theoretical structural model arises from the optimising behaviour of the borrower. This means that constraints are imposed on the relationship between the type of termination (default or prepayment) and the underlying state variables (interest rates and house prices).

This results in structural models that produce mortgage values and termination behaviour that is different from what is actually observed in practice. Recall that borrowers can choose to prepay their mortgage in full, at any time prior to the maturity of the loan if interest rate changes make this financially favourable to do so but not all borrowers do so and similarly with default.

In purely option-theoretic mortgage valuation, prepayment is modelled with an overarching and interacting American - style call option. Standard option pricing models using this approach determine that if current interest rates decrease sufficiently (so that it is financially favourable) then borrowers should prepay, the instant the value of the mortgage to the lender is greater than the cost of prepaying, in order to avail themselves of the re-finance opportunity of a new mortgage. The term '*Rational Prepayment'* to describe this result: by minimising their mortgage costs in this way Borrowers optimise their behaviour.

In practice, as found empirically, borrowers do not behave in such a ruthless way. Some prepay their mortgages later than when the model indicates they should - some mortgages are not prepaid even when their contract rate is above current mortgage rates that may indicate that borrowers act as though they face transaction costs that far exceed the explicit costs of refinancing. But there could also be a great deal of '*inertia'* preventing borrowers from prepaying. By this what is meant is a lack of knowledge by the borrower of availability of a better contracts or the even existence of refinancing possibilities, possible prepayment penalties, explicit transaction costs and such like.

Borrower expectation can also be used to illustrate a lack in optimal refinancing by indicating that borrowers choose to delay prepayment if they believe that the mortgage rate will fall further. Also of note is that mortgage values themselves can exceed par (the initial loan amount) by more than the transactions costs of refinancing. Thus actual prepayment therefore appears suboptimal, relative to the optimal behaviour implied by these models.

It is thus now well known that basic mortgage valuation models are unable to replicate mortgage values greater than par (the initial loan amount). So to model prepayment by utilising an American call option, produces mortgage values that are lower than those observed in reality.

To overcome these inherent difficulties, prepayment could be modelled with a different type of option. A barrier- type option, in particular a Parisian call option. (see Sharp et al. 2008). By doing this, a lag in prepayment being exercised from when it is financially optimal can in fact be simulated. In addition it is also possible to achieve mortgage values greater than par, within the rational structural framework.

By producing this time lag endogenously, the prepayment decision is still the result of an optimising behaviour by the borrower but by considering this time lag as a representation of the difficulty a borrower has in deciding whether it is optimal to prepay at any time, the borrower incurs a cost that reduces the value of the prepayment option and so increases the mortgage value even without consideration of any explicit transaction costs involved.

7.3.1 Barrier Options

The term barrier option can be attached to a number of different options. In barrier options, a specified asset price S_b is selected as the barrier value. During the life of the option, this barrier may or may not be crossed. In a *knock-out* option the contract is forfeit, if the barrier value is crossed at any time during the whole life; in *knock-in* options the contract is activated only if the barrier is crossed.

The barrier S_b may be set above or below the current asset price, S_0: If $S_b > S_0$ we have an 'Up' option; if $S_b < S_0$ we have a 'Down' option. These features are then combined with the payoffs of call and put options to define a range of barrier options. An *occupation-time derivative* is an example of an exotic barrier option where the action on crossing the barrier is not as straightforward as that described above.

The specification for these occupation-time derivatives includes *consecutive,* which refers to the number of prescribed consecutive time steps for the knock-in / knock-out feature to be activated, once the underlying is beyond the barrier. If the underlying returns back across the barrier the clock is reset to zero. For *cumulative*, reference is made to a summation of the time steps for every excursion across the barrier.

Consecutive options are also known as Parisian options, and cumulative options are known as ParAsian options. In principle, the barrier might be monitored continuously; but in practice, only periodic monitoring may be applied.

Both analytic pricing formulae and numerical pricing methods are available for certain barrier options and the pricing of barrier options continues to be the subject of research in its own right.

Finite difference schemes have also been developed that use direct discretisation of the governing PDE. The approach to be adopted here is that extensively discussed in Sharp et al. (2008) and outlined below.

In this approach the underlying, monitored to check whether the barrier is crossed, is the mortgage value. This is dependent upon two stochastic processes and gives rise to a free boundary, which must be set so that an appropriate scheme can be applied in the state space beyond the barrier. Beyond the barrier when the barrier clock is activated, a separate governing PDE is to be solved.

7.3.2 Valuation Framework

The general valuation framework is the same as for the standard traditional model detailed previously in Chapter 3. The main difference is that a new model regarding borrower behaviour now governs the prepayment option. As given previously, the mortgage contract consists of the value of the remaining payments to the lender $A(r,t)$ and the borrowers embedded options to Default and Prepay. The default option $D(H,r,t)$ is modelled as a monthly European Put option, where the borrower decision is modelled using financially ruthless default behaviour.

The Prepayment option $P(H,r,t)$ has a new assumption that dictates when this option is exercised. Insurance components $I(H,r,t)$ and $CI(H,r,t)$ only pay out if the borrower defaults as previously discussed. Rather than prepaying as soon as the mortgage value is equal to the total debt, some borrowers wait until the mortgage value is equal to or greater than the total debt for the decision time $\bar{T}$ and then prepay the mortgage.

Prepayment is now modelled as a Parisian call option rather than an American call option. The American free boundary, on one side of which prepayment occurs and on the other it does not is replaced with a different free boundary. Instead of the underlying asset being monitored across a barrier, the value of the mortgage is monitored above a time dependent barrier- the total debt TD(t). The barrier position in V corresponds to the free boundary location in (H, r) state space. At each instant in time the barrier TD(t) for V, is known but the free boundary location that corresponds to the mortgage value being equal to the barrier is not.

These values of H and r are required to value V, so that the modified PDE can be valued in the appropriate region of (H, r) state space. The free boundary must be positioned such that $V(H,r_b,t,\bar{t}) = TD(t)$ where r_b is the interest rate value that separates the prepayment region from the continuation region and the Total debt TD(t) is given by $TD(t) = [1 + c(T - t)]OB(i)$.

Notice that now the value of the mortgage and the other mortgage components are now functions of an additional time variable, the time until prepayment occurs.

Since it is required that the time the mortgage value spends above the prepayment barrier is recorded, there is a necessity to introduce a new time variable – the time until the decision to prepay is made call this time $\bar{t}$

All the components which are affected by the borrower terminating the contract prior to its maturity now include this variable, however the value of the remaining payments does not actually depend upon the actions of the borrower and so this is not a function of $\bar{t}$

To recap, when the value of the mortgage to the lender $V(H,r,t,\bar{t})$ is less than the prepayment barrier the usual valuation PDE applies. When the mortgage value increases such that it becomes equal to or greater than this barrier, then the new valuation PDE will apply since the waiting time until knock out begins to decrease. The effect is to produce an additional derivative in the PDE with respect to time, in order to track the time until the prepayment decision. The PDE to value a general mortgage component $F(H,r,t,\bar{t})$ that is affected by the borrower's decision time is thus

$$\frac{1}{2}H^2\sigma_H^2\frac{\partial^2 F}{\partial H^2}+\rho H\sqrt{r}\sigma_H\sigma_r\frac{\partial^2 F}{\partial H\,\partial r}+\frac{1}{2}r\sigma_r^2\frac{\partial^2 F}{\partial r^2}+\gamma(\theta-r)\frac{\partial F}{\partial r}+(r-\phi)H\frac{\partial F}{\partial H}+\frac{\partial F}{\partial t}+\frac{\partial F}{\partial \bar{t}}-rV=0$$

7.3.3 Payment Date Conditions

Now that the new PDE and the new time variable have been defined new payment-date conditions for the mortgage components need to be stated to close the problem.

At Maturity

At maturity the option to prepay is not relevant. The payment-date conditions for the value of the remaining payments, the mortgage's value and the value of default and insurance components are unchanged. These conditions thus remain the same as for the standard UK FRM Model and are as given previously.

At Earlier Payment dates

Here, there is the possibility of prepayment and so the payment-date conditions change from those modelled using the optimal call conditions of the standard UK FRM model.

Value of the remaining payments

This is not affected by either of the borrower's options to terminate and so the condition for this component remains unchanged.

Value of the other mortgage components

For the value of the mortgage to the lender we have

$$V^-(H,r,t,\bar{t}) = \begin{cases} \min[(V^+(H,r,t)+MP),H] & \text{if } \bar{t} > 0 \\ TD(t) & \text{if } \bar{t} = 0 \end{cases}$$

when the time until prepayment is zero, prepayment occurs and the mortgage is worth the value of the total debt at this time.

Default occurs when the value of the House is so low that the borrower decides not to make the scheduled monthly payment. The option to default is worthless when prepayment occurs, this occurs when the value of the mortgage is greater than the total debt for the decision time corresponding to $\bar{t} = 0$.

In this situation the option components take on the following values

$$D^{-}(H,r,t,\bar{t}) = \begin{cases} D^{+}(H,r,t) & \text{if } V^{-}(H,r,t) = V^{+}(H,r,t) + MP \text{ and } \bar{t} > 0 \quad \text{No default} \\ A^{-}(r,t) - H & \text{if } V^{-}(H,r,t) = H \text{ and } \bar{t} > 0 \quad \text{In default} \\ 0 & \bar{t} = 0 \end{cases}$$

If default occurs then prepayment option cannot be exercised giving

$$P^{-}(H,r,t,\bar{t}) = \begin{cases} P^{+}(H,r,t) & \text{if } V^{-}(H,r,t) = V^{+}(H,r,t) + MP \text{ and } \bar{t} > 0 \quad \text{No default} \\ 0 & \text{if } V^{-}(H,r,t) = H \text{ and } \bar{t} > 0 \quad \text{In default} \\ A(r,t) - TD(t) & \bar{t} = 0 \end{cases}$$

When prepayment occurs, option values are calculated by the rearrangement of V=A-D-P to determine P, the option to default is worthless and the value of the mortgage is the total debt.

Insurance perspective

For the insurance asset the option components take on the following form

$$I^{-}(H,r,t,\bar{t}) = \begin{cases} I^{+}(H,r,t) & \text{if } V^{-}(H,r,t) = V^{+}(H,r,t) + MP \text{ and } \bar{t} > 0 \quad \text{No default} \\ \min(\lambda[TD^{-}(t) - H], \Delta) & \text{if } V^{-}(H,r,t) = H \text{ and } \bar{t} > 0 \quad \text{In default} \\ 0 & \bar{t} = 0 \end{cases}$$

The occurrence of prepayment makes the value of default worthless so that the Insurance is also valueless.

7.3.4 Implementing the Numerical Solution

As previously identified two distinct PDE’s require their solutions to be determined under the new prepayment model.

These must both be solved numerically. The usual transformation in the time dimension is made for both time variables. The solution method involves backwards solution from maturity so $\tau = T - t$ (t is the time until prepayment date in the month) and the corresponding transformation for the borrower decision time is $\bar{\tau} = \bar{T} - \bar{t}$ ($\bar{t}$ being the time until the prepayment decision by the borrower).

Defining the lattice structure

For the numerical solution the equally spaced grid of points is defined as before in the H, r, and t dimensions with the new time dimension being represented as follows: Let the grid spacing in the new time dimension be p and let the new dimension be subdivided into m intervals then $\bar{t}_m = mp$. By definition p is the same as the grid spacing, s, in the t dimension.

As before i and j are the number of nodes in the House Price and Interest Rate dimensions respectively and n and m are the number of time steps dividing each month of the contract and the borrower decision time respectively.

Discretising the governing PDE's

The value of asset F after transformation onto the unit square, represented by W will then have an approximation that is represented by

$U_{i,j}^{n}$ for each (i, j, n) triple below the barrier, as previously discussed and

$U_{i,j}^{n,m}$ for each (i, j, n, m) quartet above the barrier.

Below the barrier: The governing valuation PDE is discretised as discussed before.

Above the barrier: The modified governing PDE is discretised in an exactly similar manner for the transformed variables x and y, but with the two time dimensions being taken care of the in the following manner where this time

$W(x, y, \tau, \bar{\tau}) = F(H(x), r(y), t(\tau), \bar{t}(\bar{\tau}))$

and $\frac{\partial W}{\partial \tau} \cong \frac{U_{i,j}^{n+1,m} - U_{i,j}^{n,m}}{s}$; and $\frac{\partial W}{\partial \bar{\tau}} \cong \frac{U_{i,j}^{n+1,m+1} - U_{i,j}^{n+1,m}}{p}$

This then leads to the modified valuation PDE to be represented as follows with

$a(x, y, t) = \frac{1}{2}\sigma_H^2 (H(x))^2 \omega^2 x^4$;
$b(x, y, t) = \frac{1}{2}\sigma_r^2 r(y) \varphi^2 y^4$;
$c(x, y, t) = \rho \sigma_H \sigma_r H(x) \sqrt{r(y)} \varphi \omega x^2 y^2$;
$d(x, y, t) = \sigma_H^2 (H(x))^2 \omega^2 x^3 - (r(y) - \phi) \omega x^2 H(x)$ and
$e(x, y, t) = \sigma_r^2 r(y) \varphi^2 y^3 - \gamma(\theta - r(y)) \varphi y^2$ as defined previously

$$a(x,y,t)\left[\frac{U_{i+1,j}^{n,m} - 2U_{i,j}^{n,m} + U_{i-1,j}^{n,m}}{h^2}\right] + b(x,y,t)\left[\frac{U_{i,j+1}^{n,m} - 2U_{i,j}^{n,m} + U_{i,j-1}^{n,m}}{k^2}\right] +$$
$$c(x,y,t)\left[\frac{U_{i+1,j+1}^{n,m} - U_{i+1,j-1}^{n,m} - U_{i-1,j+1}^{n,m} + U_{i-1,j-1}^{n,m}}{4hk}\right] + d(x,y,t)\left[\frac{U_{i+1,j}^{n,m} - U_{i-1,j}^{n,m}}{2h}\right] +$$
$$e(x,y,t)\left[\frac{U_{i,j+1}^{n,m} - U_{i,j-1}^{n,m}}{2k}\right] - \left[\frac{U_{i,j}^{n+1,m} - U_{i,j}^{n,m}}{s}\right] - \left[\frac{U_{i,j}^{n+1,m+1} - U_{i,j}^{n+1,m}}{p}\right] - r(y)U_{i,j}^{n,m} = 0$$

Rearranging this gives the expression for the modified PDE:

$$sU_{i,j}^{n+1,m+1} - (s-p)U_{i,j}^{n+1,m} = [p(1-sr(y)) - \frac{2asp}{h^2} - \frac{2bsp}{k^2}]U_{i,j}^{n,m} + \left(\frac{bsp}{k^2} + \frac{esp}{2k}\right)U_{i,j+1}^{n,m} + \left(\frac{bsp}{k^2} - \frac{esp}{2k}\right)U_{i,j-1}^{n,m} + \left(\frac{asp}{h^2} + \frac{dsp}{2h}\right)U_{i+1,j}^{n,m} + \left(\frac{asp}{h^2} - \frac{dsp}{2h}\right)U_{i-1,j}^{n,m} + \frac{csp}{4hk}\left(U_{i+1,j+1}^{n,m} - U_{i+1,j-1}^{n,m} - U_{i-1,j+1}^{n,m} + U_{i-1,j-1}^{n,m}\right)$$

Dealing with the free boundary

The prepayment barrier level introduces a free boundary problem that divides the *(H, r)* state space into two regions one below the boundary where the usual PDE holds and one above the boundary where the modified PDE holds.

Above the prepayment boundary the mortgage value is calculated with the modified PDE where borrowers are assumed to prepay after decision time $\bar{T}$ has elapsed. The free boundary problem therefore occurs at each time step. Problematically the free boundary will not in general coincide with the regular spaced nodes of the lattice and so a discrete approximation of this boundary also needs to be found. This may be determined using an iterative boundary tracking process as used by Sharp et al. (2008) where

At each time step:

1: Take an initial guess as to the free boundary location, let this be $\Omega_j^{old}(i)$

2: Value $U_{i,j}^{n,m}$ for all *i* and for j=0 to j=$\Omega_j^{old}(i)$ using the modified PDE and for j= $\Omega_j^{old}(i)+1$ to j= max(j) using the standard PDE

3: Determine the new location $\Omega_j^{new}(i)$ from the condition $V > TD$ where the value is greater than the prepayment barrier

4: If $\Omega_j^{new}(i) = \Omega_j^{old}(i)$ then stop since we have found the approximate position at the current time step.

5: If $\Omega_j^{new}(i) \neq \Omega_j^{old}(i)$ then let $\Omega_j^{old}(i) = \Omega_j^{new}(i)$ and then repeat from 2: (the new guess to the free boundary is the approximate location from the previous time step)

In this way as we pass across the H dimension a maximum value of interest rate *r* for each *H* is found. The modified PDE provides the mortgage value for interest rates less than the approximate free boundary and house prices greater than the approximate free boundary, with the usual PDE providing the mortgage value elsewhere.

7.3.5 Simulation Results

In order to present and discuss the numerical results provided by this model, the basic set of economic parameter specifications have been previously identified in Table 2 of Chapter 6. The choice of these fixed parameters for the results in this section is made in accordance with these base parameters. The parameters which were allowed to vary are the initial interest rate r(0) along with the house-price volatility σ_H and the interest-rate volatility σ_r.

Table 17: Comparison of Mortgage Asset components values for different prepayment assumptions.

(25yr Repayment Mortgage with AFEE and ETP) (£)

r(0)	Decision Time	Volatility	Mortgage Value V		Default D	
		House Price	ν=0.05	ν=0.10	ν=0.05	ν=0.10
		Int. Rate				
6%	0	σ=0.05	88640	87895	30	800
		σ=0.10	88640	88300	45	515
	T/2	σ=0.05	88650	87860	15	808
		σ=0.10	88644	88290	27	665
	T	σ=0.05	88654	87790	10	852
		σ=0.10	88640	88240	36	829
	3T/2	σ=0.05	88632	87745	22	891
		σ=0.10	88635	88120	40	919
r(0)	Decision Time	Volatility	Prepayment P		Insurance I	
		House Price	ν=0.05	ν=0.10	ν=0.05	ν=0.10
		Int. Rate				
6%	0	σ=0.05	1360	1220	20	797
		σ=0.10	6480	5980	12	300
	T/2	σ=0.05	1255	1196	10	812
		σ=0.10	6305	5800	6	360
	T	σ=0.05	1115	1159	8	848
		σ=0.10	6200	5600	8	412
	3T/2	σ=0.05	1040	1118	9	894
		σ=0.10	6084	5428	10	518

r(0)	Decision Time	Volatility / House Price / Int. Rate	Mortgage Value V ν=0.05	Mortgage Value V ν=0.10	Default D ν=0.05	Default D ν=0.10
8%	0	σ=0.05	88655	88070	25	710
		σ=0.10	88648	88280	20	640
	T/2	σ=0.05	88660	88090	25	749
		σ=0.10	88648	88390	15	680
	T	σ=0.05	88660	88059	24	790
		σ=0.10	88648	88339	22	835
	3T/2	σ=0.05	88660	88028	24	814
		σ=0.10	88648	88267	27	888

r(0)	Decision Time	Volatility / House Price / Int. Rate	Prepayment P ν=0.05	Prepayment P ν=0.10	Insurance I ν=0.05	Insurance I ν=0.10
8%	0	σ=0.05	3230	2910	9	520
		σ=0.10	9550	8802	5	228
	T/2	σ=0.05	3165	2787	7	545
		σ=0.10	9340	8727	4	260
	T	σ=0.05	3090	2687	8	580
		σ=0.10	9157	8360	4	300
	3T/2	σ=0.05	3020	2642	8	618
		σ=0.10	8954	8240	4	365

r(0)	Decision Time	Volatility / House Price / Int. Rate	Mortgage Value V ν=0.05	Mortgage Value V ν=0.10	Default D ν=0.05	Default D ν=0.10
10%	0	σ=0.05	94220	88250	15	700
		σ=0.10	88650	88347	10	630
	T/2	σ=0.05	94118	88280	14	720
		σ=0.10	88650	88442	7	680
	T	σ=0.05	94023	88254	14	768
		σ=0.10	88650	88420	11	824
	3T/2	σ=0.05	93895	88221	14	759
		σ=0.10	88650	88376	12	862

r(0)	Decision Time	Volatility / House Price / Int. Rate	Prepayment P ν=0.05	Prepayment P ν=0.10	Insurance I ν=0.05	Insurance I ν=0.10
10%	0	σ=0.05	5578	5160	4	352
		σ=0.10	12800	12222	1	170
	T/2	σ=0.05	5480	5005	5	385
		σ=0.10	12540	11874	1	190
	T	σ=0.05	5380	4872	4	384
		σ=0.10	12286	11459	1	210
	3T/2	σ=0.05	5250	4790	4	425
		σ=0.10	12024	11278	1	260

The three panels of Table 17 above show the associated mortgage component values for four prepayment assumptions: ‘**0**’ is the simple prepayment assumption (decision time T = 0) and this was the basis for all the results of Chapter 3 on mortgage valuation using ruthless prepayment. Results are also shown using the new prepayment model for the decision times T = 0.5 month (**T/2**), T = 1 one month (**T**) and T = 1.5 months (**3T/2**).

To understand the results in the table, we bear in mind the equilibrium condition and recall that the value of the mortgage is the value of the remaining payments minus the borrower’s options to terminate the contract.

As the decision time T increases it is expected that the value of prepayment will decrease, as borrowers are less inclined to prepay, and as the decision time does not affect the default decision, the only way that the equilibrium condition is satisfied will be if the value of payments decreases; this occurs if the contract rate decreases. The limiting cases of allowing the decision time to tend to zero results in the simple ruthless exercise assumption and of allowing the decision time to tend to infinity causes the prepayment region to disappear, illustrating that prepayment would never occur. These features are evident in the table for the different initial interest rates and volatilities tested, in particular as the decision time increases the equilibrium contract rate would be expected to decrease (and as a result the value of payments decrease), and the value of prepayment decreases.

The mortgage values in the table do not vary significantly, for fixed initial interest rate r(0) and fixed volatilities σ_H and σ_r as the decision time T increases. This is because the majority of the effect of the decision time is focused in the prepayment region of the state space which is away from the mortgage value vital to the equilibrium condition (at the initial house price and the initial interest rate).

The following observations are general features of mortgage valuation independent of the decision time analysis:

(i) As the initial interest rate r(0) increases, the possible prepayment region grows and as the prepayment region is for rates less than r(0) (otherwise a contract in equilibrium would not exist), this means that there is a greater possibility of prepayment. Given this scenario the value of prepayment increases as r(0) increases, and the required equilibrium contract rate would also increase as a result.

(ii) For corresponding parameters between the table levels, as house-price volatility σ_H increases it is expected that the contract rate decreases. Although the value of default increases as σ_H increases (by analogy to the result that an increase in stock price volatility raises the value of a stock option), it may be expected that this would increase the contract rate but as insurance covers default, and since if default occurs prepayment cannot (meaning that prepayment decreases in value as σ_H increases), means the required contract rate would actually fall as a result of an increase in house price volatility.

(iii) For corresponding parameters between the table levels, as interest-rate volatility σ_r increases it is expected that the contract rate increases. This is more clear-cut in that, as σ_r increases then the value of prepayment increases; this would potentially augment the borrower's position, unless the equilibrium contract rate also increased.

7.3.6 Concluding Remarks

The new model of prepayment identified by Sharp et al. (2008) and examined in this section provides a economical structural means of modelling a borrower's termination behaviour that appears `*irrational*' according to the results of a basic optimal exercise model and illustrates that results can be obtained that are outside the extent of simple rational models. By incorporating barrier type option into the mortgage valuation framework a more advanced (compared with the simple ruthless approach to prepayment modelling) borrower decision process can be employed, where a rational exercise structure is retained in a modified form.

The results here confirm those of Sharp et al. (2008) by showing that it is possible to achieve mortgage values above par (the initial loan amount) within a structural model by including the borrower's decision time before prepayment is made. The results show that the direct effect of increasing the decision time is to increase the value of the mortgage above par *inside* the prepayment region (under a simple option-theoretic model this would not be possible); correspondingly the value of prepayment decreases.

This approach of creating a lag in prepayment is a method of modelling borrowers who do not exercise their option to prepay when it appears financially optimal to do so. By varying the decision time borrowers who do not act in the same way the heterogeneity of borrowers can be modelled under this framework. In addition embedded options within the pricing framework of mortgage backed securities - whose value is determined by the behaviour of a large group of individuals who cannot be counted on to act according to a simple rational model – may be valued.

7.4 Incorporating Current Economic Conditions

This section is solely intended to provide insight into a potential mechanism for including in the traditional model of Chapter 3 a factor relating directly to the state of the current economic climate and represents an avenue essentially for future research activity.

In order to incorporate directly the current economic conditions into the standard model of Chapter 3 the following modification may be undertaken.

The standard model is modified by a '*discounting*' device such that the value of the mortgage is

$$V(H, r, \tau(i)) = [A(r, \tau(i)) - D(H, r, \tau(i)) - P(H, r, \tau(i))] \cdot \lambda(H, r, \tau(i))$$

Where $0 \leq \lambda(H, r, \tau(i)) \leq 1$

and as before $\tau(i)$ denotes the i^{th} payment date.

Now λ can be decomposed into three components:

$$\lambda = \lambda_{IND} \cdot \lambda_{\text{HPI}} \cdot \lambda_{\text{SCO}}$$

where λ_{IND} is tied to the conditions of the economy, λ_{HPI} is tied to the liquidity of the mortgage market and λ_{SCO} is tied to the quality of the mortgage. These discounting factors being defined on the interval $(\tau(i\text{-}1), \tau(i)]$. Adding such a discount factor as λ to the model may make it more suitable to the present turbulent economic situation.

Economic Conditions: λ_{IND}

The UK Index of Leading Economic Indicators (UK LEI Index) is designed to predict the economy's direction. This index is a composite of a select group of economic statistics that are known to swing up or down well in advance of the rest of the economy. UK Leading economic indicators are time series that peak before the UK macro-economy peaks and reach a trough before the UK macro-economy moves into an expansionary phase and are generally used as a means of identifying the early stages of a recession or a recovery.

This type of indicator is compiled and released on a monthly basis for different countries by private business research groups like the Conference Board, Academetrics and also by the International Community via the OECD for the expressed purposes of judging direction in the economy (and housing market in the short-run), these series are available on a frequent basis and are considered to provide early signals of future economic activity. These leading indicators are based on business cycle stylised facts offering additional judgmental information that is likely to affect sentiment in the housing market.

The factor λ_{IND} could be constructed in the following way:

Let Δ(LEI%) be the reported change of the LEI Index in the month prior to the payment date that is $\tau(i\text{-}1)$, then we can define

If Δ(LEI%) is positive then $\lambda_{IND}(\tau(i)) = 1$

If Δ(LEI%) is negative and the absolute Δ(LEI%) < 1 then

$\lambda_{IND}(\tau(i)) = 1 - (\Delta(\text{LEI\%}) /100)$

If Δ(LEI%) is negative and the absolute Δ(LEI%) ≥ 1 then $\lambda_{IND}(\tau(i)) = 0$.

Mortgage Market Liquidity: λ_{HPI}

The Financial Times House Price Index FT_HPI is a reasonable candidate for this type of assessment since the current market is assessed for home sales including an expectation of future trends. The FT_HPI is produced monthly, based upon Land Registry Data and by providing an additional degree of completeness (through its forecasts) it provides a better illustration of the underlying trends in prices for the purpose of inclusion as a discount factor here.

Other possibilities - that can be considered - include the Halifax HPI, but it suffers the disadvantage of being based upon approvals rather than completions (sales) likewise the Nationwide HPI suffers in relation to its more biased construction – Nationwide data is specific to their own mortgage approvals.

So if the factor FT_HPI denotes the value of this index as reported on the month prior to the payment date then:

Let Δ(FT_HPI%) be the reported change of the FT_HPI Index in the month prior to the payment date that is $\tau(i\text{-}1)$, then we can define

If Δ(FT_HPI%) is positive then $\lambda_{HPI}(\tau(i)) = 1$

If Δ(FT_HPI%) is negative and the absolute Δ(FT_HPI%) < 1 then

$\lambda_{HPI}(\tau(i)) = 1 - (\Delta(\text{FT\_HPI\%}) /100)$

If Δ(FT_HPI%) is negative and the absolute Δ(FT_HPI%) ≥ 1 then

$\lambda_{HPI}(\tau(i)) = 0$

Mortgage Contract Quality: λ_{SCO}

For use in this situation is a discount factor is chosen so as to be a constant for the whole mortgage term. The likelihood that a loan will not be repaid and it will fall into default is calculated by financial institutions for each borrower and can be determined for groups of borrowers with substantially similar characteristics. The Credit Score of the borrower is taken into account in determining its value. This is a simple and straightforward approach for financial institutions having access to credit bureau information from Experian or Equifax. Failing this their own application score information derived from in-house score card at contract origination could be used. At mortgage origination we can then define $\lambda_{\text{SCO}}(0)$ to be equal to the borrower's probability of default as evidenced by their credit score / application score, thus on the interval $(\tau(i\text{-}1), \tau(i)]$ we may define $\lambda_{SCO}(\tau(i)) = \lambda_{\text{SCO}}(0)$.

7.4.1 Illustrative application of the model

As a trial experiment to provide an initial assessment - investigate how well the model replicates actual mortgage values - a sample of 50 mortgage values at origination (originations span the 12 month period of 2003) were taken at random from a dataset supplied by a financial institution who wishes to remain anonymous and mortgage value under these economic adjustment modifications determined.

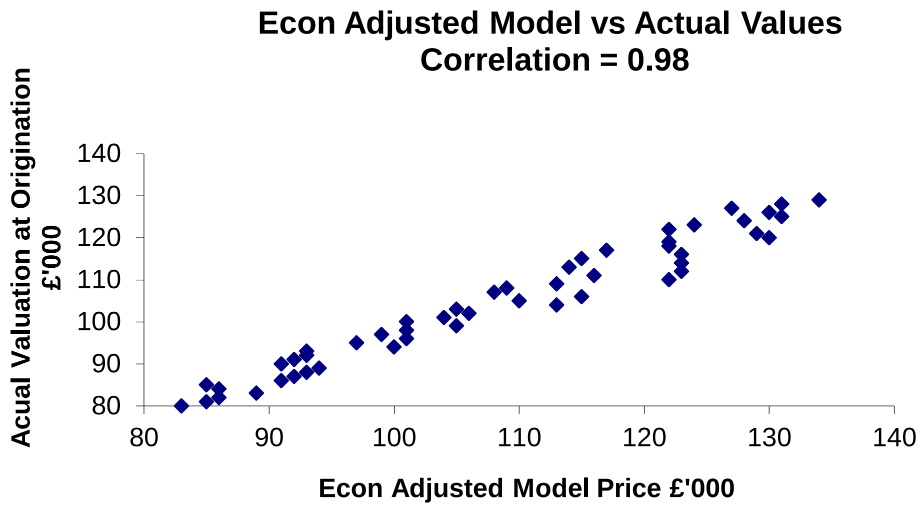
Time series data for the UK Index of Leading Economic Indicators (UK LEI Index) was compiled with some difficulty manually from published Conference Board bulletins online and that for the Financial Times House Price Index FT_HPI was also similarly manually compiled from published Academetrics bulletins also online. The Application Scores for the sample of mortgages were each converted into a probability of default by utilising a simple and easy to perform logistic transformation since the actual probability of default employed by the financial institution is not known.

The observed mortgage prices, indices and probability of default used in the pricing analysis are presented in an Appendix F

7.4.2 Results of the Model Application

As illustrated in Figure 18 below, there is a significant linear relationship between model and observed mortgage values for this random sample.

Figure 18: Discounted Model versus Actual Mortgage Value at Origination



The model is meant to replicate the actual prices observed, however where the modelled and actual prices differ, it is unclear whether the model has simply mispriced them or has a bias. Overall however the model does seem to fit the observed prices well. The estimated correlation coefficient between the model and actual prices is 0.98, indicating a strong linear relationship between the model and the observed values.

7.5 Modelling in Turbulent Times: An Extended Structural Approach

Existing credit default theory tends not to provide a direct link between the causes and effects of residential mortgage termination. It thus gives the appearance that it is unable to evaluate satisfactorily the credit risks inherent in a turbulent market environment, as was recently experienced in the mortgage credit market crisis. This global financial crisis exposed the shortcomings of current credit risk models.

To overcome these weaknesses it is suggested in this section that a framework that links causes and effects of mortgage termination should be considered in an attempt to construct quality models that have a high predictive ability and are not too heavily dependent upon empirical data.

The focus of this section is to use the fundamental Black Scholes Merton option pricing theory to model the value of the mortgage, the option to default and the probability to default. From within such a framework a credit default theory in which the probability of default PD and the loss given default LGD are determined endogenously and consistently may be developed.

7.5.1 Background Context and Other Preliminaries

Using Basel II credit risk principles for the measurement of expected loss (PD, LGD, EAD) Joseph (2007) demonstrated how a mortgage pricing model can be divided into a number of key component parts. These constituents include the embedded options (default and prepayment) of the mortgage along with all their associated transaction costs. Through industry best practice his work demonstrates by example how PD's may be estimated with what is termed the *traditional 3C's of credit underwriting* long time used by prime lenders:

Character: Borrower willingness to repay their mortgage loan as measured through an application score or behavioural score.

Capacity: Borrower ability to repay their mortgage loan obligation as measured by a model of affordability.

Collateral: Borrower honest appraisal of the asset value of their property as measured through a structural model of credit risk.

It is the aim of this section to propose that credit default risk may initially be modelled jointly by – the latter two of these 3C's – the two fundamental concepts underlying practically all definitions of credit default, through the processes of Delinquency (modelled by a stochastic random variable related to borrower liquidity) and Insolvency (modelled by a stochastic random variable related to property value).

In this way the macroeconomic causes of turbulent financial conditions may be linked to the microeconomic effects of change in personal circumstance of a borrower – or even to the corporate environment – that may lead to the possibility of default.

The literature reviewed in Chapter 2 provides an indication that most studies that examine the impact of mortgage default risk on the valuation of mortgages focus on asset values as the sole default trigger, assuming that borrowers are ruthless and will default when the property value falls below the mortgage value. The typical conclusion from these studies being that the *current* LTV will be a key, if not the primary driver, of default risk.

Other theoretical and empirical research, however, suggests that asset value is not the sole default trigger, and find that 'ability to pay' variables are somewhat more important determinants of default.

Such research endeavour consistently supports the importance of borrower equity; however, the evidence on the significance of borrower ability to pay through an affordability model involving the payment-to-income ratio in explaining default incidence is also brought into focus and it is also worth noting from this work that the probability of default is relatively higher if a rational borrower is under pressure to exercise their in-the-money option and insufficient cash flow exists.

This suggests that it is useful to view default as driven primarily by the level of equity, but with the default condition modified when borrowers have an inability to meet mortgage payments. Of course either or both may be present when a borrower chooses to default

Mortgage regulation in 2004 has also put greater emphasis on lenders to demonstrate responsible lending. This has lead to what is now the norm that each lender provides evidence of an individual borrower's ability to repay. In this way lenders are seen to be treating their customers fairly and the borrower is seen to demonstrate a genuine ability to service their mortgage loan obligation

To properly reflect a rational borrower's default decision, it is reasonable for models therefore to include both an equity explanation through a link to property value and an ability to pay explanation through a link to borrower income as triggers of default.

7.5.2 Credit Default Concepts

Two concepts, Delinquency and Insolvency underlie the majority of definitions of credit default. *Delinquency* defined as a failure to meet a loan payment by a due date. *Insolvency* defined as a situation where assets are less than liabilities.

The majority of uses of the term '*credit default*' in reduced form models really encapsulate the concept of delinquency.

The Basel II definition of default is essentially a definition of delinquency whereas structural models really make use of the concept of insolvency.

With the assumption that insufficient cash flow to service the debt obligation constitutes the primary cause of a credit default then the situation is that if an individual borrower is unable to meet the mortgage payment obligation by the due date then that borrower becomes delinquent.

If this cash flow problem continues unresolved within a given prescribed period of time, then the lender assumes that the borrower may be insolvent, and the mortgage loan is considered to be in default. The theoretical process through which a borrower proceeds to loan termination may then be encapsulated in the following:

The borrower has insufficient cash flow to service the loan (delinquency) triggers an evaluation of the net equity position of the borrower (a solvency assessment) which may lead to a possible conclusion of insolvency (position of negative equity) causing the loan to be terminated - the lender expects to incur losses.

To mitigate these losses the lender then initiates some sort of proceedings to recover the secured asset upon which the loan is made. Therefore in reality a credit default for a loan secured by property (Mortgage) is a sequence in time of two separate circumstances: delinquency followed by insolvency.

For a mortgage we can assume delinquency is a necessary requirement for default, since a borrower could by other means refinance the loan, from say positive equity or other net assets and prevent default. In general for such a loan, both delinquency and insolvency are assumed both necessary and sufficient for credit default.

7.5.3 A Capacity Measure: Debt Serviceability (Loan affordability) Method

There are two methods that can be used for the assessment of the debt servicing capacity of a borrower. First, there is the traditional method of the *Debt Service Ratio* which compares borrower debt (including other fixed payments) to the borrower gross income from employment and investment. This method allows lenders to impose a general 'rule of thumb' limit of (⅓ rd) on the portion of gross income that could be prudently used to cover debt payments

Second, there is the *Net Income Surplus method.* Here the borrower is to have a minimum surplus of net (after-tax) income taking into account debt servicing, other fixed payments and living expenses. Notice that this approach uses after tax income and not gross income. Explicit account is also taken of living expenses and debt servicing requirements other than the mortgage loan. Lenders can use this method to examine if a borrower could continue to meet their mortgage obligation under adverse conditions. This assessment may be undertaken systematically; for instance in discounting the income stream, adding a margin to living expenses or changing interest rates.

Under the net income surplus method, a maximum loan amount is calculated on the basis of 'eligible' sources of income, an estimate of basic living expenses, the interest rate and any discount margins applied to these components. Net income surplus method also utilises a ratio measure.

The numerator of the measure is the after tax income of the borrower (the net disposable income) less all debt servicing obligations other than the mortgage loan (personal, car, credit card, HP and other forms of extended credit) and living expenses (estimated through household poverty index + margin) and family size.

The denominator of the measure is the sum of interest and principal repayments on the housing loan, estimated using current interest rates, the loan amount and a 25yr notional maturity.

This measure is expected to be in excess of one. The more the measure exceeds one, the greater is the cushion available to meet loan obligations without the need to draw upon sources other than income or to compromise the level of household consumption.

Lending amounts based upon a ‘rule of thumb’ for net income coverage measure are typically between 1.5 and 2.5 times the debt repayment (suggesting that most borrowers are well insulated from potential problems). A net income coverage cushion lower than this indicates to a lender that the borrower could be more vulnerable to increases in interest rates or declines of income (not necessarily through job loss but also from reduction in overtime or bonuses) as the effects of a downturn in the economy filter through.

Modelling Delinquency

Delinquency occurs when a borrower chooses or is unable to make the scheduled loan payment by the date due, caused by liquidity failure. A liquidity failure is when there is insufficient disposable income after allowing for cost of living and other expenses, to meet debt obligations (the mortgage payment).

The inherent risk comes from the fact that loan serviceability changes over time, due to changes in borrowers individual circumstances coupled with changes in the economic environment.

For instance, a loan which could have started off as being easily serviceable may develop into a struggle for the borrower due to unanticipated adverse developments

We introduce a liquidity model which estimates the probability that the stochastic delinquency variable x_L falls below unity, where

$$\begin{aligned} x_L &= \frac{NetDisposableIncome}{LoanPayment} \\ &= \frac{AfterTaxIncome - LivingCosts - OtherExpenses}{MortgagePayment} \\ &= \frac{AfterTaxIncome - LivingCosts - OtherExpenses}{MortgagePaymentRate \times LoanAmount} \end{aligned}$$

In mathematical terms $x_L = \dfrac{W - X - D}{\left(rL \big/ (1-(1+r)^{-n}) \right)}$

where in the numerator W = after tax income, X = living expenses and D = other expenses (debt repayments) and in the denominator L = loan amount, r = interest rate and n = original term of the loan. Note that other forms of Debt Coverage Ratio used in the literature do not include the cost of living expense as here.

Let *t* be continuous and denote duration time, that is time since the beginning of the loan. Then the Loan Serviceability Ratio (LSR) variable given an initial value $x_{L,}$(t=0), can evolve a-priori in any stochastic manner since the underlying variables on which the LSR depends can themselves evolve according to a number of stochastic processes.

Once the stochastic process controlling the evolution of the delinquency variable is specified, we can determine its probability distribution at a later time, and from this calculate the probability for delinquency where $x_L(t) < 1$.

For the moment with simplicity in mind assume a Gaussian stochastic process for x_L(t). With this assumption (and others) the LSR evolves according to the well known diffusion equation of Black Scholes (1973) for option pricing.

The solution of this is given by the lognormal distribution for LSR with two parameters: the drift rate μ_L and volatility σ_L so that after time t this evolution of LSR can be described by $z_L(t) = \frac{\ln(x_L(t)) + (\mu_L - \frac{1}{2}\sigma_L^2)t}{\sigma_L \sqrt{t}}$

and x_L (t) is a standard normal variate with mean zero and unit variance.

We may term z_L(t) the ‘distance to liquidity failure’, refer to Brandimarte (2006: pp 98-100) for an outline of the standard derivation.

Let P_L (t) denote the probability that the borrower will face a cash flow problem in meeting a loan payment by time *t*. Then the probability of delinquency or liquidity failure when LSR < 1is given by

P_L (t) = Φ (-z_L(t))

where Φ *(.)* is the standard normal (cumulative) probability function.

7.5.4 A Collateral Measure: Credit Risk Model

There are two recognised approaches to credit risk models*: reduced form* models and *structural* models (Crook and Bellotti: 2009). Most credit risk models in use are reduced form models which are usually a linear subset of a more general econometric model.

Structural models originated from Merton (1973) who applied equilibrium theory of option pricing to corporate debt. Strictly speaking his model is not a credit default theory, but rather a theory of risk premia determination based upon the assumption that traders eliminate arbitrage opportunities in a market at equilibrium.

Essentially the Merton model estimates the probability that the stochastic insolvency variable $$x_E = \frac{Assets}{Liabilities}$$

falling below unity, indicating the borrower has negative equity. Given a number of assumptions including that the stochastic process to be a standard random walk of a Gaussian process we obtain the closed form solution to Black Scholes option pricing formulae.

The basic idea of insolvency as the cause of default in the Merton model may however not be sufficient enough to explain all cases of default; other reasons occur. For instance, there may be individual borrowers who, whilst technically insolvent, are still able to ward off default because they continue to make the necessary debt repayments. In other words, liquidity or the ability to pay (service the loan commitment) may also be a decisive factor in the actual occurrence of default.

Modelling Insolvency

In the original Merton Model for insolvency (Merton 1974) the random variable that determines credit default risk is the assets to liabilities ratio that defines a situation of negative equity if it is less than one. For residential mortgages the corresponding random variable is the property value-to-loan ratio VLR or the inverse of the LTV ratio that is:

$$x_E(t) = \frac{V_t}{L_t}$$

Note that both V and L vary with time t, V as a result of property price changes and L as a result of loan amortisation.

Again, by modelling $x_E(t)$ as a Gaussian stochastic process we can describe its evolution with a lognormal distribution. The evolution of VLR after time t can then also be described by

$$z_E(t) = \frac{\ln(x_E(t)) + (\mu_E - \frac{1}{2}\sigma_E^2)t}{\sigma_E \sqrt{t}}$$

where the drift rate is μ_E and σ_E denotes volatility. For comparison, in the Merton Model the drift rate is replaced by a riskless interest rate due to the absence of arbitrage at market equilibrium. We may again refer to $z_E(t)$ as the 'distance to negative equity'. Refer to Brandimarte (2006: pp 98-100) for an outline of the standard derivation.

Let $P_E(t)$ be the probability that the borrower will face a negative equity position by time *t*. Then the probability of negative equity when the VLR < 1 is then given by

$$P_E(t) = \Phi(-z_E(t))$$

where $\Phi(.)$ is the standard normal (cumulative) probability function.

7.5.5 Modelling Probability of Default, Expected Loss and Loss Given Default

Let borrower default by time t be given by D with default being determined solely by liquidity failure L and negative equity E. Note that L and E may be separated by a time lag.

Now given that credit defaults are determined jointly by the two random variables $z_L(t)$ for liquidity failure and $z_E(t)$ for negative equity, credit default may be modelled by the random variable $z_D(t)$ where $z_D(t) = f(z_L(t), z_E(t), t)$. In general *f* is described by the solution of a stochastic partial differential equation in these two stochastic variables subject to certain assumptions about the underlying processes as for example in the work of Kau, Keenan and Kim (1993, 1994)

In the case of mortgage loans secured on property the two stochastic variables for delinquency and insolvency may be correlated a-priori. Under the assumptions of Gaussian processes for $x_L(t)$ and $x_E(t)$ the probability of Default $P_D(t)$ is determined by a Bivariate Normal probability density function with a given correlation coefficient ρ.

That is $P_D(t) = P(z_L(t), z_E(t), t) = \dfrac{Exp(-\frac{1}{2}Q)}{2\pi\sqrt{1-\rho^2}}$

with *Q* being given as

$$Q = \frac{z_L^2(t) - 2\rho z_L(t) z_E(t) + z_E^2(t)}{1-\rho^2}.$$

Refer to Rice (1995: pp 79-83) for details. (See Appendix G for a derivation)

Now in the special case where $\rho = 0$ the twin causes of default are independent and uncorrelated so the probability of default is then given simply by

$$P_D(t) = P_L(t)P_E(t) = \Phi(-z_L(t))\Phi(-z_E(t))$$

With reasonable selected values for set of four parameters: $(\mu_L, \sigma_L); (\mu_E, \sigma_E)$ it may be possible to reproduce the main observed features of a property market over a given time frame.

Similar models may be applied to any property market provided we can model the factors affecting the cash flow situation of the borrower.

Expected Loss and Loss Given Default

The probability of negative equity is P(x_E(t)<1) = Φ (-z_E(t)) where Φ (.) is the standard Normal (cumulative) probability function. The random value-to-loan ratio x_E(t) evolves subsequently according to a log normal distribution with the particular distribution depending upon the drift and volatility parameters and time.

Following Brandimarte (2006: pp 99) the value to loan ratio $x_E(t)$ at time t, evolves as

$$x_E(t) = x_E(t=0)\exp\left[(\mu_E - \tfrac{1}{2}\sigma_E^2)t + \sigma_E \varepsilon \sqrt{t}\right] \qquad \text{where } \varepsilon \sim N(0, 1).$$

Now consider the borrower as the holder of an American option. That is, a PUT option that is exercised at any time *t* provided the option has value. If the property value is less than the loan value there will be negative equity and the option will be in-the-money.

The expected value of the put option is given by the expected value of the 'pay-off' per unit of the loan value: Pay-off (x_E(t)) = max (0, $1-x_E$ (t)). By integration the expected payoff (or the gain) for the borrower V_{BG} in exercising the option is found.

Now the expected gain to the borrower V_{BG} is the expected loss to the lender V_{LL}. So we have:

$$V_{BG} = V_{LL} = \Phi(-z_E(t)) - x_E(t=0)\exp(\mu_E t)\Phi(-z'_E(t))$$

where $z'_E(t) = z_E(t) + \sigma_E\sqrt{t}$. (refer to Hull (2006: pp 310-312) for details)

Under Basel II Expected Loss, EL equates to the Probability of Default PD times Loss Given Default LGD times Exposure At Default EAD, Now if as a result of negative equity default occurs, the expected value of the loss given default V_{LGD} will be simply given by the following expression assuming exposure at default (EAD) equals 1:

Expected Loss = Loss Given Default times Probability of default so

$$V_{LGD} = \frac{V_{LL}}{\Phi(-z_E(t))}$$ (See Appendix G for a derivation)

7.5.6 General Numerical Simulation

Using simulation to illustrate the application of the model, a simple assessment of residential mortgage default risk using this model is outlined for a household with a net disposable income of £8 500 per annum with a mortgage on a property of value £100,000.

For the delinquency variable we assume the loan serviceability ratio LSR

$$x_L = \frac{AfterTaxIncome - LivingCosts - OtherExpenses}{MortgagePayment}.$$

For the insolvency variable we assume the reciprocal of the loan to value ratio LTV given by equation $x_E = \frac{V}{L}$. The loan approval process captures the relevant data from the borrower to provide estimated of LSR and VLR at origination for each loan. Both x_L and x_E are variable over time.

Evidence suggests that macroeconomic factors have a potential impact on the universe of housing loans. (Crook and Banasik: 2005) so given microeconomic and macroeconomic assumptions about how wages, inflation, consumer credit usage, interest rates and property prices are likely to change in the future period, it may be possible to estimate the 4 model parameters $(\mu_L, \sigma_L); (\mu_E, \sigma_E)$ to predict how LSR and VLR will evolve over time.

From their time dependent probability distributions we can calculate for any given time ahead the probability of default, loss given default and expected loss for any given loan. Here we show in principle how macroeconomic factors can be introduced into the model for Probability of default.

Given previously $x_L = \frac{W - X - D}{\left(rL \big/ (1-(1+r)^{-n}) \right)}$ we may write $U = W - X - D$ and $m = \frac{r}{\left(1-(1+r)^{-n}\right)}$ then $\ln(x_L) = \ln(U \big/ mL)$ where U = Net disposable income that is available to service the mortgage loan and m = loan payment rate. Now it can be shown that $\mu_L = \mu_U - \mu_m$ and $\sigma_L^2 = \sigma_U^2 + \sigma_m^2$

Where μ_U = average rate of change of the average wages and salaries in the economy, σ_U = log volatility of average wages and salaries in the economy, μ_m = average rate of change of interest rates in the economy and σ_m = log volatility of interest rates in the economy.

Applying similar considerations to $x_E = \frac{V}{L}$, for home loans, the rate of change of equity random variable $\mu_E = \mu_p + r_a$ and $\sigma_E^2 = \sigma_p^2$ Where we have μ_p = the rate of change of property prices, r_a = the loan amortisation rate and σ_p = the log volatility of house prices.

Parameter estimates, as indicted in the tables below, may be obtained from data sources giving actual changes in average wage, inflation, mortgage interest rates and property prices over a specified period. For instance Council of Mortgage Lenders (CML) and Office for National Statistics (ONS) here in the UK or for the USA, the Bureau of Economic Analysis (BEA)

By calculation, prediction of what would happen to housing loans if the assumed environmental conditions remain constant over the forecast period can be made.

For the purpose of the illustration hypothetical values are

Table 18: *In Period of Rising Property Prices: Boom Condition*

	Drift rate	Volatility
x_E	15%	5%
x_L	20%	15%

Assumption: Prices rise rapidly and employment and interest rate environments are mildly positive

Table 19: *In Period of Declining Property Prices: Bust Condition*

	Drift rate	Volatility
x_E	-10%	25%
x_L	0%	25%

Assumption: Moderate and steady price declines with neutral employment and interest rate environments but a rather increased volatility in the risk factors

Using the results of this simulation with an initial value of LSR = 1.2 and various LVR levels corresponding to a range of different LTV ratios 80% to 95% the model can be used to assess how the features of the loan change over time for the two economic situations.

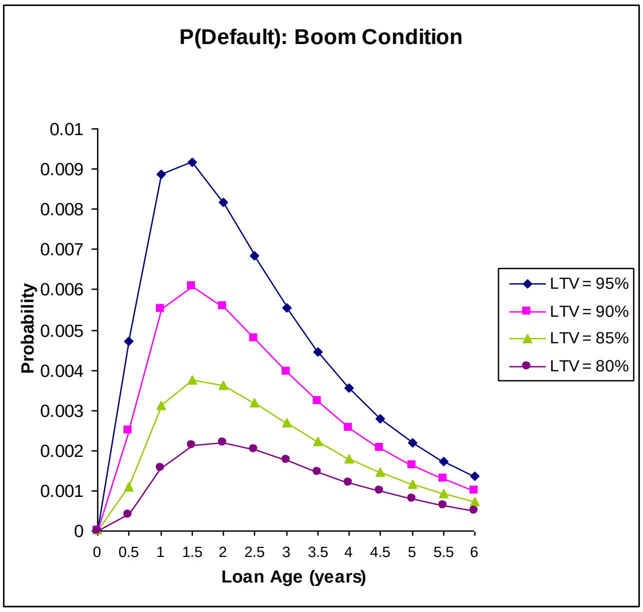
Using the hypothetical values for a steadily rising and a steadily falling property market as indicated in tables 18 and 19 a simple interpretation is presented.

Probability of default

The figures shown below are derived from the values in Table 18 (Hypothetical Boom Conditions) and Table 19 (Hypothetical Bust conditions) and equation for (Probability of Default)

Figures 19 and 20 present the results for several loan-to-value LTV ratios, the probabilities of default over time at yearly intervals, from origination for a period of 6 years.

Figure 19: P (Default): Boom Conditions



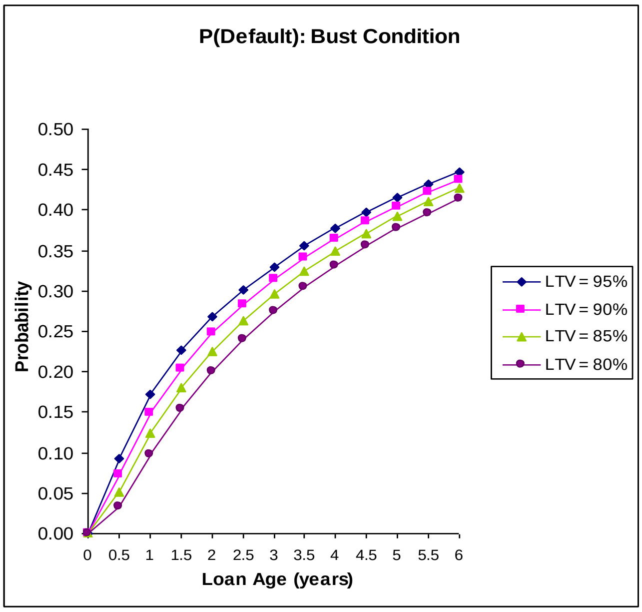


Figure 20: P (Default): Bust Conditions

From these two charts we can see that in a steadily rising property market, the probability of default for the residential loan peaks around 18 months into the loan. (Fig.19)

There is quite a dramatic increase in the probability of default as the steadily rising property market changes to a falling property market. (Fig. 20)

Loss Given Default

The figures shown below are derived from the figures in Table 18 (Hypothetical Boom Conditions) and Table 19 (Hypothetical Bust conditions) and equation for (Loss given Default)

From these two charts it can be seen that LGD proportion is variable increasing with time since uncertainty increases with time.

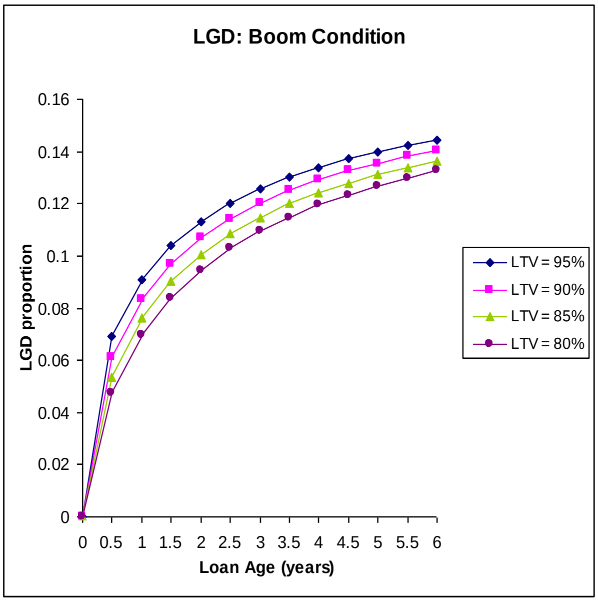


Figure 21: LGD: Boom Conditions

For high LVR loans the LGD proportion quickly rises above 10% in just a couple of years with potential consequences for claims against lenders insurance. (Fig.21).

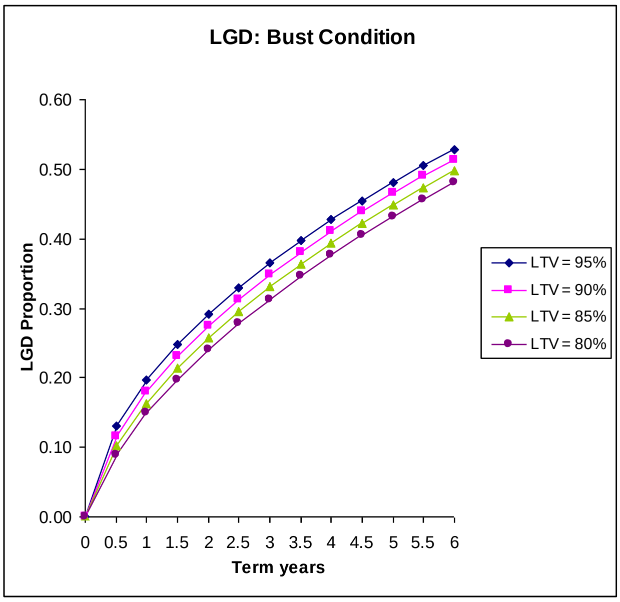


Figure 22: LGD: Bust Conditions

In a falling property market the corresponding values are several times larger and there appears to be a more pronounced increase in LGD proportion over time. (Fig.22)

These findings appear contrary with the assumption of a constant LGD in current credit default models. Suggestive in Figure.21 is the feature of recent rising property market conditions where the LGD tends to plateau after several years.

It is useful to note how the probability of default and LGD are correlated is dependent upon external conditions and the lenders individual loan portfolio.

7.5.7 Model Results: Empirical Application to UK Data

Now consider a further simple application of the model, to UK data from the sources identified previously. For the UK we have the following:

UK Housing market Booms and Busts:

Housing Market Price rises in years 1983-1989 and 1996-2007 (limit of data used)

Housing Market Price falls in years 1990-1995 give rise to the following parameter values summarised in the tables below

Table 20: *In a typical Period of Rising UK Property Prices: Boom Condition*

	Drift rate	Volatility
x_E	11%	35%
x_L	5%	20%

Assumption: Prices rise relatively quickly and employment and interest rate environments are mildly positive.

Table 21: *In a typical Period of Declining UK Property Prices: Bust Condition*

	Drift rate	Volatility
x_E	-3%	59%
x_L	6%	23%

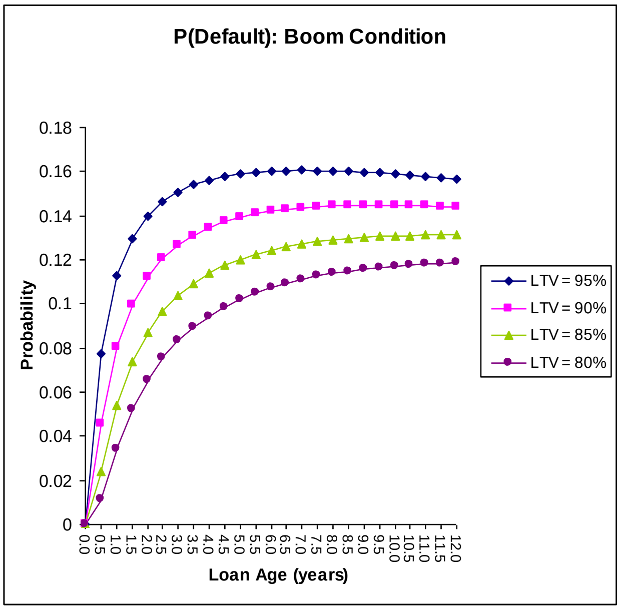
Assumption: Moderate and steady price declines with neutral employment and interest rate environments but a rather increased volatility in house prices.

The volatility parameter is calculated by the method of Hull (2006:pp287-288)

Probability of default

The figures shown below are derived from the values in Table 20 (Typical UK Boom Conditions) and Table 21 (Typical UK Bust conditions) and equation for (Probability of Default)

Figure 23: P(Default): UK Boom Conditions



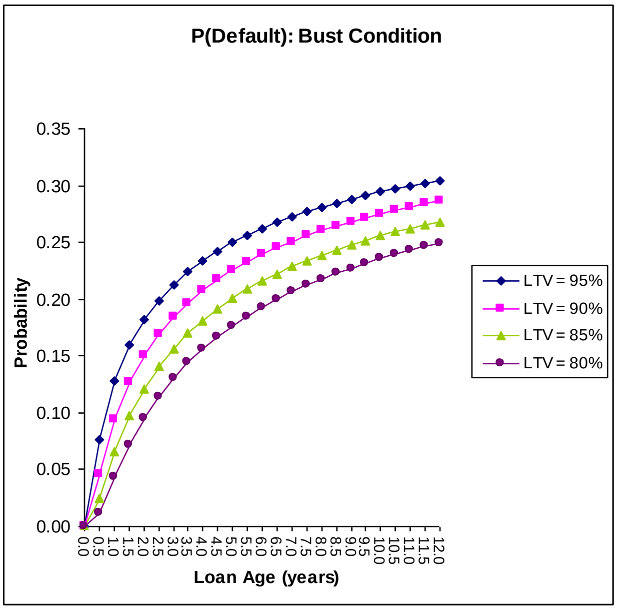


Figure 24: P(Default): UK Bust Conditions

Under the UK market conditions summarised by tables 20 and 21, the Figures 23 and 24 present the probabilities of default cumulatively over time at yearly intervals, from origination for a period of 12 years. This is done for several loan-to-value LTV ratios.

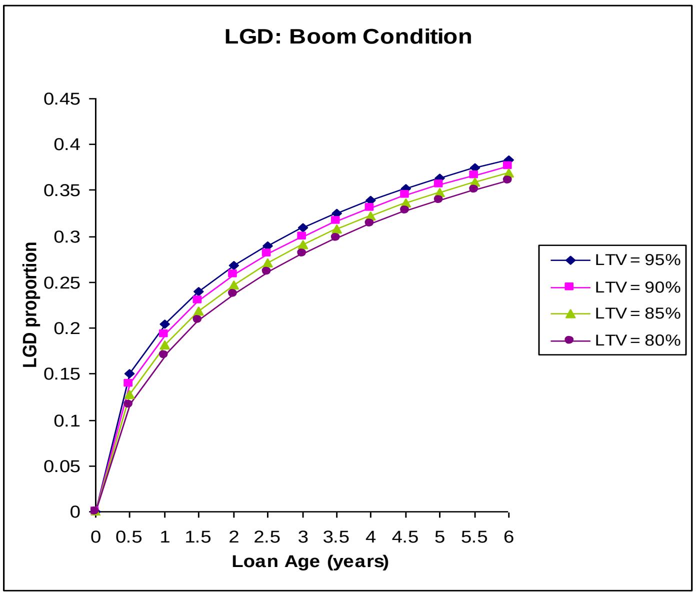
Both figures demonstrate that the higher the LTV the higher the probability of default. Figure.23 suggests that the probability of default is greatest within the first 5 years and then tapers slowly away. Due in part to the accumulated equity build up in the house under the favourable market conditions of steadily rising house prices.

Figure 24, shows a rising probability of default over the entire period under the less favourable market condition of steadily falling house prices.

These results show some consistency with the work of Kau, Keenan and Kim, (1994) *Loss Given Default*

The figures shown below are derived from the figures in Table 20 (Typical UK Boom Conditions) and Table 21 (Typical UK Bust conditions) and equation for (Loss given Default)

Figure 25: LGD: UK Boom Conditions



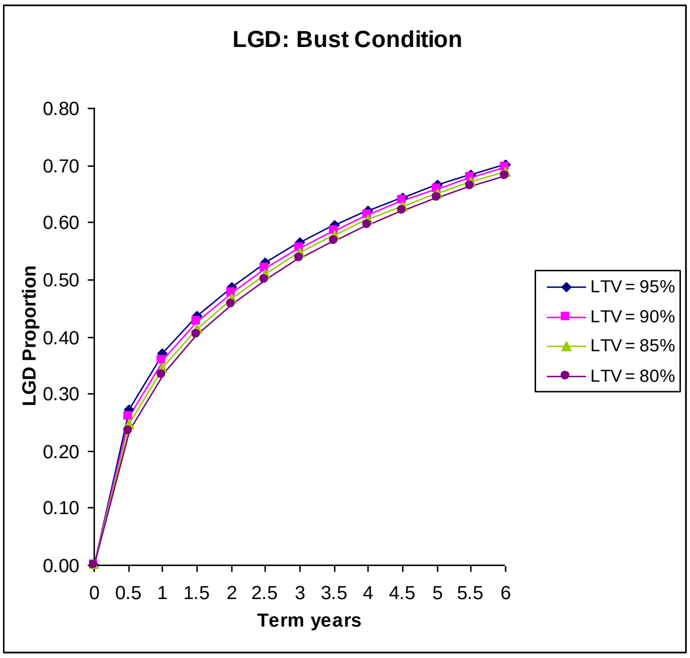


Figure 26: LGD: UK Bust Conditions

From these two charts it can be seen that LGD proportion is variable and increases over time since uncertainty increases with time. These figures are broadly similar to the hypothetical case previously considered.

For high LVR loans the LGD proportion quickly rises above 25% in just a couple of years with potential consequences for claims against lenders insurance. (Fig.25). In a falling property market the corresponding values are at least twice as large and there appears to be a slightly greater increase in LGD proportion over time. (Fig.26)

7.5.8 Concluding Comments and Further Research

This section has outlined a proposal for a framework where the causes and effects of credit default can be investigated systematically. Critical to this framework is a methodology that prime lenders have been utilising for a long time – the so called 3C's of credit. By consideration of two of these C's, the approach taken is that credit default results from both delinquency and insolvency.

A theoretically correct methodology for application to fixed and variable rate mortgages has been developed. A definition of credit default has been provided that enabled through the use of spreadsheet simulation an illustration of how the model works giving an appreciation of the practical application of the model. Further, application to typical UK Data furnishes results which are found to consistent with previous research studies in the United States.

Enhancing this framework to include a measure of the credit character of the borrower through a credit scoring model, measuring the other competing risk of early redemption of the mortgage by the borrower through modelling their propensity to prepay and the incorporation of the effects of income protection insurance for the borrower and on mortgage insurance for the lender would represent avenues for further exploration.

Chapter 8

Empirical Modelling

8.1 Introduction

This chapter presents the data analysis. It addresses systematically the structure of influences affecting the default and prepayment options in mortgage contracts from an empirical standpoint. Subsequent sections introduce some terminology, notation and assumptions that will serve as a basis for the construction of empirical models and their implementation.

The rest of the chapter is outlined follows: Section 8.2 opens the discussion by defining the default criteria used and Section 8.3 provides a discussion of the risk characteristics and features that make up the model specifications within subsequent chapters. Each of the remaining three sections 8.4 to 8.6 the methodology and implementation of the different empirical models is described. In section 8.4 the modelling focus is upon time varying covariates. In section 8.5 the competing risks of mortgage termination are modelled. The chapter is completed by section 8.6 where models focus upon unobserved heterogeneity and correlation over time.

8.2 Default Risk: Sources and Definition

This section and the one that follows, opens with a discussion of the definition of risk for the supplied event history data, along with descriptions of the sources of this risk and its nature along with other supplemental data to be used in the empirical analysis.

8.2.1 Defining Risk for Event Histories

To examine the extent to which the definition of default may impact the empirical nature of the competing risk of Prepayment and Default, as well as the degree to which unobserved heterogeneity may be identified, models are estimated with extended definitions of default to include those loans that are 90+ days delinquent as opposed to those loans exclusively in the process of foreclosure or indeed actually foreclosed.

The extension of the definition of default reflects the borrower's view of default with respect to default exercise, whilst the restricted definition being more closely related to default outcome, reflects the lenders view.

8.2.2 Definition of Default

We define default as the first event representing foreclosure of the borrower's interest in the property asset – onset of the foreclosure process.

That is, the default event is represented as the first instance of a 90-day delinquency regardless of ultimate disposition. Foreclosure would thus be defined as being 90 days past due and proceedings initiated.

8.3 Model Specification: Risk Characteristics and Risk Features

Informed by the body of literature previously reviewed, those characteristics identified and considered as significant to a mortgage modelling exercise are determined and their formulation and potential for inclusion are noted for use in the subsequent analyses.

Now the prepayment and default behaviour of each loan in the pool of loans under investigation in this research is to be forecast from those characteristics that may influence the borrowers' likelihood of early mortgage termination. Hence, the influence of each characteristic on prepayment and default is discussed.

8.3.1 Equity Risk Features

Starting with the *Equity Risk Features* detailed as follows:

- Loan to Value (LTV) Ratio

LTV is the ratio between the principal balance on the mortgage and the appraised value of the property serving as security for the loan itself. There are two forms of LTV input used in default modelling

a) The Original LTV: that is, the LTV at origination of the loan and;

b) The Current LTV: the calculated LTV for each month post-origination for the duration of the observation period.

The LTV is calculated by taking the outstanding balance of each loan and dividing it by an estimate of the current market value of the property each month. The current market value of the property is calculated by taking the given original property valuation and adjusting.

Property valuations are updated by indexing their given valuation according to a House Price Index. The index that is used is that of the Nationwide House Price Index. In addition, due to the geographic distribution of the data house price adjustments are calculated on a coarse UK region dependent basis i.e. separately for properties in England, Scotland and Wales.

The **OrigLTV** variable may be considered as a proxy for the credit worthiness of the borrower. An estimate of the current loan to value ratio **CurrentLTV** on a mortgage holder's property can be used to represent the extent to which the Default (put) option is 'in-the-money'. Higher CLTV ratios are therefore associated with an increase in the likelihood of default, this being attributable to the progressively smaller proportion of equity that the borrower has in the property.

To include both the current LTV and the original LTV in the same specification, may give a high degree of correlation between the two measures and may be problematic. To get around this potential problem should it arise, it may be addressed by including the initial size of the mortgage (log initial loan balance) and the value of the property at the origination of the mortgage (log initial house value) as independent explanatory variables.

OrigLTV contains the value of the loan to value ratio at mortgage loan origination, and is further categorised using the variable **LTVcat**, into the following bands:

OrigLTV <=75%, LTVcat = 1
75%< OrigLTV <=80%, LTVcat = 2
80%< OrigLTV <=85%, LTVcat = 3
85%< OrigLTV <=90%, LTVcat = 4
90%< OrigLTV <=95%, LTVcat = 5
OrigLTV>95% , LTVcat = 6

- Loan Size

Given that mortgage lenders generally limit the maximum loan size based upon income multiples, larger loans are only available to borrowers with higher incomes.

Since higher incomes may be subject to greater volatility in the event of an economic downturn, it is considered that larger loans are more risky than smaller ones. For example, a high income borrower who may rely on significant bonuses may find it more difficult to maintain their financial status when forced into a new job position as a result of a changing economic environment. Whilst larger loans might therefore reflect a good credit rating they also bear higher servicing costs and this in turn may lead to an increased likelihood of default. Likewise low purchase price for a property may reflect wealth factors such as less absolute value of collateral for borrowing and household liquidity constraints.

The original loan balance **loansize** is included to control for the borrowers financial flexibility and preference for debt, It is expected that larger loans are more likely to prepay, given the same refinancing incentive.

- Equity

Equity is the difference between the value of the property and the amount of loan secured against it. The smaller the equity, the smaller the potential financial benefit the borrower can retain from the property, and the lower the incentive to maintain the scheduled monthly mortgage payments, so that declines in property value increase the likelihood of default - by increasing the chances of negative equity.

Consequently the research reported here addresses this issue by a determination of the amount of equity the borrower has in their property, the EquityAmt and then setting an associated dummy variable **Neq_ind** equal to 1 for those observations where equity is negative, 0 otherwise.

Note that the variable Neq_ind which is a time varying covariate is introduced to denote the presence of negative equity. Additionally the following variable is also determined. The probability that a given mortgage loan will have an 'in-the-money' default option in a given observation period - after Deng et al. (1996) - is calculated representing the Likelihood of Negative Equity,

NegEquity = $N\left(\frac{\log(pvbal)-\log(mktval)}{\sqrt{\sigma^2}}\right)$

This is the probability that any given property value is below the mortgage balance at each monthly observation, here *pvbal*: is the present value of the remaining mortgage payments at the current market interest rate and *mktval*: is the current market value of the property estimated by using the index of house price growth and its volatility.

8.3.2 Ability-to-Pay Risk Features:

The borrowers' predisposition to default is clearly related to the ability to make timely mortgage repayments on an on-going basis. Affordability and other income and employment related aspects are very important, particularly at a time of rising consumer indebtedness in the UK giving rise to serious concerns as to the ability of people to repay their debt should favourable economic and housing conditions deteriorate.

- Single Income versus Joint Income

When scheduled mortgage payments are serviced by two separate incomes, if one income becomes unavailable, being able to rely on a co-borrowers income mitigates the likelihood of default. The variable **inc_ind** an associated dummy variable is set equal to 1 for those mortgages where there is only one income, 0 otherwise. *As such the payments on a mortgage serviced by a single income are considered by this research to be more at risk than those that are serviced by the 2 separate incomes*.

- Loan to Income (DCR) Ratio

DCR is a measure of loan affordability and is commonly used by lenders to determine how much they are prepared to advance on a mortgage. DCR is calculated by dividing the loan balance by the total income for the household. Many lenders also use more sophisticated affordability measures to take into account other financial commitments.

In a recent report the CML indicated that approximately 85% of UK lenders continue to use DCR measures[18], either as the sole determinant of maximum loan amount or as a cross check to other affordability measures. Although it is likely that more complex affordability measures are better indicators of risk than the simple DCR, the components of these measures are not consistent across lenders. *As such it is considered in this research that DCR is a simple but effective means of assessing affordability.*

[18] Council of Mortgage Lenders, "UK Mortgage Underwriting", April 2006

In the past, loans were generally restricted to 3 times single income or 2.5 times the joint income. Recently these income multiples have been generally allowed to increase by lenders to 4 times earnings for borrowers with low LTV ratio or good credit history. An income multiple variable, **IncMult,** is therefore included. Higher DCR's are generally seen as a sign of greater financial commitment and may make a borrower much more susceptible to default in the case of life changing experiences – so called 'trigger events' – such as Divorce or Unemployment. As a consequence *it is considered appropriate in this research that loans based on DCR's that exceed 3.5 times are more at risk*

- Self Certification

Self certification is used by borrowers who want to obtain a mortgage without having to demonstrate their earnings to a standard required by conventional mortgage underwriting criteria. They simply declare their income without any supporting documentation like payslips. Typically these borrowers are self-employed; commission based or contract workers and those with incomes from a variety of sources. Taken as a whole, such borrowers show a large degree of variability in income over time. There are thus additional risks associated with self-certified mortgages coupled with the common concept that a self-certified borrower could not afford the mortgage loan under normal lending criteria.

Lenders endeavour to offset this risk by a variety of means, the most common being:

i. A plausibility check to ensure that the stated job type fits within a reasonable salary range.

ii. More conservative credit scoring and lower LTV ratios so as to deter borrowers from taking out a mortgage that they cannot afford.

Despite these additional safeguards, the higher level of arrears experienced with this type of product means that *it is considered in this research that self certified loans tend to be more risky than benchmark loans.*

The dummy variable **selfcerts** is set equal to 1 if the loan is to an individual who self certifies their income, 0 otherwise.

- Self Employment

Self-employed borrowers who do not self certify their income need to provide the mortgage lender with documentary evidence of their earnings. However compared to borrowers who are employees, self-employed borrowers tend to have lower stability in terms of monthly income.

In addition self-employed borrowers often need to undertake large financial investments in order to set up their own businesses that may make them more vulnerable in an increasingly stressful financial environment. *Hence it is considered in this research that having a self-employment status tends to be more risky than the normal employment status.*

Occupational Status, through the variable **Advnc_occ_cde1** contains the occupational status categories of the Mortgage borrower where:

Advnc_occ_cde1 = 1 if Employed
Advnc_occ_cde1 = 2 if Self-Employed
Advnc_occ_cde1 = 3 if Unemployed

8.3.3 Mortgage Contract Design and Initial Choices

Risk features inherent to the that are considered important include

- Loan Product Choice

A brief summary of the most common mortgage product types is described below:

i. *Standard Variable Rate*, SVR. This is set by the individual lender and usually increases or decreases in line with the Bank of England BoE, base rate. Hence mortgage interest payments based on SVR are likely to rise or fall every time the BoE modifies the base rate.

ii. *Tracker* loan products track the rate set by the BoE plus or minus a differential for a short term (2- to 5-years) and then switch to the SVR for the remaining life of the loan.

iii. *Discount* loan products pay interest on the basis of the SVR minus a discount for a short term (2- to 5-years) and then switch to the SVR for the remaining life of the loan.

iv. *Fixed Rate* loan products pay interest based on a fixed interest rate typically below the SVR for a short term (2- to 5-years) and then switch to the SVR for the remaining term.

It is considered in this research that Tracker, Discount and Fixed rate types attract a greater degree of risk than the SVR, for these mortgages do not generally track the on-going changes in interest rates over time but are subject to a potentially more significant payment shock: An increase in payments once the mortgage product reverts to the SVR if the interest rate has increased during the 2- to 5-year short term period.

This factor also controls for any relative popularity of fixed rate to variable rate mortgages. The variable for product grouping **typ** and **FRM** contains the mortgage type information

‘Fixed’, typ=1, ‘SVR’, typ=2, ‘Tracker’, typ=3, ‘Disc’, typ=4 and ‘Base’, typ=5.

The dummy FRM = 1 if typ = 1, 0 otherwise. This factor also controls for any relative popularity of fixed rate to variable rate mortgages.

- Loan Repayment Type

As identified earlier in the thesis there are two main mortgage repayment methods currently available in the UK. Repayment and Interest Only IO. There are many variations of each of these two types. In a standard repayment type mortgage, both interest and some capital borrowed are paid back over time to ensure that the mortgage is totally paid off by the end of the Term. In contrast, IO mortgages only require the repayment of the interest on the initial principal balance until maturity, when the borrower repays the principal balance. There is a general trend toward the growing use of IO mortgages that allow borrowers to defer the payment of principal.

For example, those borrowers whose income is fairly low but who have expectations for extra financial income (from bonuses say) benefit from smaller regular payments of interest and a more flexible approach to principal repayment. In other words they enjoy a smaller short-term monthly commitment than a regular repayment borrower.

With growing levels of unsecured consumer indebtedness combined with high house prices, borrowers are likely to consider IO mortgage loans as a way to afford property that they may not be able to afford with a regular repayment scheme.

As such there is concern that IO borrowers are more likely to have stretched financial circumstances. In addition there are also concerns around the borrowers' ability to pay back the entire balance due on the mortgage at the maturity date.

The Financial Services Authority recently highlighted, that many IO borrowers did not have a strategy in place for repaying capital. Although a borrower can refinance at maturity, the market environment at that future date is unknown and as such the borrower exposes themselves to refinance risk.

- Loan Purpose

Borrowers apply for mortgages primarily for two reasons, Home Purchase or Re-Mortgage. In recent years re-mortgage activity has boomed in the UK. Typically when a borrower re-mortgages, they use the proceeds from the re-mortgage to pay down an already existing mortgage, with the same property being used as collateral. The main motivation for this type of re-mortgage, also referred to as refinancing, is usually to take advantage of a more favourable interest rate on offer by an alternate mortgage provider.

As such a growing proportion of borrowers are raising capital from their properties, hence taking on more debt. Debt consolidation is a particular form of such equity release where one loan (the re-mortgage) is taken to pay off other debts already existing.

Being a form of increasing credit exposure, it contributes to stretch borrowers' finances, potentially compromising their ability to pay their debts and so *it is considered in this research that raising capital and debt consolidation by this* *manner is to be associated with a higher likelihood of default compared with more traditional mortgages.*

By including the dummy variable **loanuse** that indicates that the Mortgage is used for either ‘House Purchase’, where loanuse = 0 or for ‘Re-Mortgage’ (standard refinancing) where loanuse = 1 allows the model to potentially control for the effects of transactions costs, and to account for the borrowers previous mortgage decisions.

a). Right to Buy, RTB

The right to buy scheme was originally introduced in the UK in 1980. Under the scheme, council tenants of registered social landlords or housing associations can buy their own homes at a low price, because part of the rent paid over the previous years of tenancy is discounted from the full market value.

Borrowers who exercise their right to buy typically have more fragile economic backgrounds and are likely to have relied on some form of financial support in the past. *It is therefore considered in this research that loans granted on the basis of this scheme are riskier than those of a standard mortgage loan.*

The **rtbs** dummy variable indicates whether the loan was granted on the basis of this scheme, rtbs = 1 if Mortgage is a Right-to-buy loan rtbs = 0, Otherwise

b). Buy to Let, BTL

A BTL mortgage is for the purchase or re-mortgage of a residential property used for investment purposes. Here the property is rented out to tenants as opposed to direct owner occupation by the borrower. In the UK BTL Lending has seen spectacular growth in recent years and corresponds to almost 10% of all UK outstanding mortgages. *BTL mortgages are considered in this research to be riskier than traditional owner occupied mortgages.* The variable **purchrsn** indicates that property is either 'BTL' Buy-to-let or 'O-O' Owner Occupied purchrsn = 0 if 'O-O', purchrsn = 1 if 'BTL'

8.3.4 Frictions and 'Trigger' Events

Finally, certain other risk features need to be borne in mind these include certain Frictions, Borrower related exogenous '*trigger*' events and Macroeconomic Conditions. In reality borrowers do not all act at the same time and they face costs whenever they decide to prepay or default.

- Transactions Costs

In this research it is therefore reasonable to assume that:

a) Borrowers face transactions costs whenever they prepay or default.
b) Borrowers do not necessarily prepay/default immediately it becomes optimal to do so.
c) Borrowers are heterogeneous, that is different borrowers face different transactions costs.

In matter of fact both the borrower and the lender bear a number of costs associated with loan delinquency, repossession and subsequent property resale. From the perspective of the lender these payments need to be subtracted from the sale proceeds. Costs include legal fees and expenditures associated with any property maintenance the sale requires and the estate agency charges.

It is considered in this research that the lender will be exposed to fee payments of approximately 3% of the sale price. Legal and other miscellaneous fees are assumed to be fixed at £2,500 (These being fairly industry typical figures although some are defined on a sliding scale over a short time period) Since there are no direct measures of transactions costs proxies must be used. Frictions and Transactions costs are included in the model through the inclusion of the various initial charges, **Initchgs**, and early redemption penalties, **RedempFeePaid**, these are combined to give a value for the potential transactions costs faced by a borrower as a result of early repayment of the existing loan and the acquisition of a new refinancing loan.

To account for the frictions of decision making (prepayment/default), a 2 month time lag is used on all independent variables that derive their values from interest rate levels. The two month time lag allows us to account for friction in the marketplace (allows for the time required for a borrower to make a decision and for a result to transpire). It is noted that in industry practice prepayment decision lags actual loan payoffs, since most mortgage lenders require 30-60 days for application processing.

- Prepayment Alternative

Due to the risk-based pricing strategies adopted by many lenders in a benign economic period, borrowers with past credit difficulties are more likely to prepay than prime borrowers, as they will be able to ‘credit cure’ in this environment and then refinance at a lower interest rate.

To capture the dynamics of the prepayment option value as it relates to changes in interest rates a variable measuring the level of interest rates relative to the contract rate and rate volatility is used.

The variable **inmnyppay** following Deng et al, (1994) is defined as

$$\text{inmnyppay} = 1 - \frac{pvbalr}{mortbal}$$

Where *Pvbalr*: is the Present Value of the remaining mortgage payments at the current market interest rate. and *Mortbal* is the Outstanding Balance.

Positive values of inmnyppay indicate that the market rate is greater than the contract rate and accordingly the prepayment option is 'out-of-the-money'. Negative values of inmnyppay indicate that the prepayment option is 'in-the-money'.

- Borrower Related Exogenous '*Trigger*' Events

Several other macroeconomic (Trigger event) variables are included, these being selected as most likely to impact on the risk of default as an early termination mechanism. For instance, borrowers may default due to a 'trigger event' leading to a loss of income and thus to ability to pay problems.

These are captured by reference to Time Series for Unemployment Rate and Divorce Rate.

a). Unemployment

The UK Regional Unemployment Rate, **Unemp**, corresponding to the three regions of the UK previously identified is included. As unemployment increases, borrowers encounter ability to pay problems leading to higher default rates. The unemployment rate therefore proxies for the joint probability of a default 'trigger event' and its severity.

b). Divorce

The UK Divorce Rate, **UKDiv_Rate,** is included. As divorce may also lead to ability to pay problems as joint resources are separated to support two households, this sometimes promotes early redemption as assets are liquidated before being divided.

8.3.5 Macroeconomic Conditions

To assess the impact of economic conditions on mortgage performance the following variables are used:

a). Interest Rates and Interest Rate Volatility

To capture the relative position of the market interest rate with respect to contract rate and rate volatility we add a variable **IR_diff** representing the Interest rate differential between the current market rate and the mortgage contract rate.

$IR\_Diff = \frac{r_c(t) - r_f(t)}{r_f(t)}$ to represent the time varying relative interest rate spread

where

r_c = current contract interest rate at time t

r_f = conventional Fixed Term mortgage interest rate at time t proxied by Nationwide Building Society Rate.

This variable also represents the relative interest rate savings associated with switching from the current SVR mortgage to a FRM.

The variables **rate** and **mktrate** contain the values of the mortgage contract rate at origination and the current market mortgage rate and these are used to give the **Rate_sp** and **rate_spcu** where Rate_sp = rate – mktrate and rate_spcu **= (**Rate_sp)^3 proxy for the incentive to refinance and its non linearity.

Kau, Keenan, Muller and Epperson, (1993) argue that interest rate volatility has a significant impact on prepayment option value with prepayment declining as volatility increases.

An Interest rate volatility variable defined as the standard deviation of the 10 year treasury bond rate, **Vol_10yr** and the 20 year treasury bond rate **Vol_20Yr** measured over the previous 24 months is used.

b). Expectation of Future Interest Rates

As an indicator of borrower expectations concerning future interest rates we include a two measures of the term structure **Spr_Ten** and **Spr_Twenty** defined as the 10 yr Treasury Bond Rate minus the 1 yr Treasury Bond Rate and the 20 yr Treasury Bond Rate minus the 1 yr Treasury Bond Rate. Given that new mortgages are indexed to the 10yr Rate a relative increase in coupon spread indicates that prepayment is becoming more valuable.

c). House Prices and House Price Volatility

The measures of House Price appreciation and volatility are derived from Nationwide Building Society Index and are used to perform monthly adjustments. Details are indicated below:

The variable **mprpo_price** contains the purchase value of the House at Mortgage loan origination, and the variable containing the current market value of the property is **mktvalprop**. It is calculated from The Nationwide regional House Price Index, **HPI**, corresponding to the three UK regions Scotland, Wales and England (inc. Northern Ireland). The variable, **loc**, contains the regional location descriptor ‘SCOTLAND’, loc=1, ‘ENGLAND’, loc=2 and ‘WALES’, loc=3.

d). The health of the Economy

Certain factors that are correlated with the strength of the economy and, thus, may lead to higher property market turnover rates are included. These additional macroeconomic time series are the Index of retail prices **RPI**, the Financial Times all share index **FTALLSH** and house price appreciation **HPI** index.

8.3.6 Other Factors

The variable **MaxTerm** contains the Term of the mortgage loan and the variable indicating first time buyer status **ftbs** are included where ftbs = 1 if the loan is to a First time buyer, ftbs = 0 Otherwise.

Also incorporated is the number of months from origination, **Loanage** to capture the impact of Mortgage Seasoning on Prepayment probabilities and also include a quadratic term $loanAge^2$ is included to control for the non-linearity in the slope of the Hazard Function.

Third Party Originations: Since mortgage lenders incur costs to obtain mortgage loan assets, profitability depends on the duration of the stream of borrower payments. The duration of loan payments depends therefore in part on the hidden actions of the agent who may encourage the borrower to refinance.

The variable, source **src**, contains a customer/borrower type:

‘Existing’, src=1, ’Direct’, src=2, and ‘Intro’, src=3 is included to address the Agency Problem with Mortgages originated by third parties.

The variable **hichgs** if a high charging structure applies to the loan, hichgs = 1, otherwise, hichgs = 0

The variable **Advnc_mar_sta1** contains the marital status of the Mortgage borrower, where Advnc_mar_sta1 = 1 if borrower Single/Widowed, Advnc_mar_sta1 = 2 if borrower Married and Advnc_mar_sta1 = 3 if borrower Separated/Divorced

Application Scores based upon all information relevant to a loan application are most often used to determine which credit requests are clearly acceptable under established underwriting guidelines and which need further review. Hence, whilst not a behavioural score as such the variable **advnc_score_app** contains the value of the application score at mortgage loan origination, it may be considered a summary of application characteristics with higher scores representing better credit quality.

These scores are further categorised into the following bands:

advnc_score_app <=300, scorband = 1
300< advnc_score_app <=450, scorband = 2
450< advnc_score_app <=500, scorband = 3
500< advnc_score_app <=600, scorband = 4
advnc_score_app > 600, scorband = 5

To capture the seasonality effect on Borrower Mobility, a season variable is included. The impact of calendar effects upon borrower mobility is well established so a season variable accounts for the fact that more relocation takes place in spring and summer relative to the rest of the year. Quarterly dummy variables, **Q1 – Q4**, are included to control for the spring and summer months in which property sales are high.

Q1 takes the value 1 if the observed month is from March to May and 0 otherwise,

Q2 takes the value 1 if the observed month is from June to August and 0 otherwise,

Q3 takes the value 1 if the observed month is from September to November and 0 otherwise,

Q4 takes the value 1 if the observed month is from December to February and 0 otherwise.

This is now followed by a brief look at certain necessary assumptions that will serve as the basis for the construction of the empirical models and their subsequent implementation and assessment.

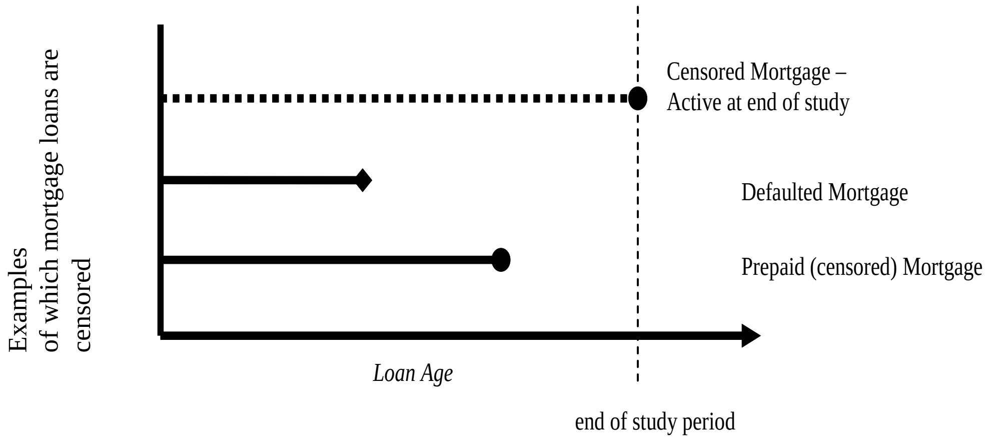
8.3.7 Censoring and Censoring Bias

The body of mortgages available to study was observed over the period Jan2000 (base year) through mid-2009 with July 2009 being the end point of the period of observation, so that right censoring is present. This is the most common form of censoring. This data presents type 1 right censoring which is where the mortgage loans' observation time reaches the end of the defined study period and the loan has not yet experienced the event of interest (early termination). Type 2 right censoring does not occur since we do not end the mortgage loans' observation period on a pre-specified number of events occurring and neither does left censoring since an early termination event has not already occurred when mortgage loans' observation time period begins.

To summarise, some mortgage loans have the early termination event of interest, early in the study period and others have the early termination event of interest later in the study period. Likewise some mortgages enter the study period after the observation period start and/or leave the study period early, but the majority do not have an early termination event of interest during the entire study period and are simple right censored at the end.

This censoring is always considered to be non-informative. Note that the supplied data set does not include any mortgages that are redeemed (matured loans) and so cease to be observed during the study.

Figure 27: Diagrammatic representation of Censoring in the Dataset



There are many mortgage loans remaining at the end of study observation period, that have neither prepaid, defaulted nor have been completely amortised.

Note that from the perspective of Prepayment (Default): A Default (Prepayment) in effect constitutes a censored observation seeing that Prepayment (Default) decision is interrupted.

The survival time is measured from the month the mortgage was originated. If a mortgage never experiences an early termination event during the observation period then they are censored and the observation time is the period of time from when the mortgage originated to the end of the study period – indicating the last time they were observed.

Thus the variable **Surv_Time** contains the time until a mortgage experiences an early termination event or is censored and the censor variable **Delta,** indicates whether or not an early termination event occurred (Delta = 1 indicates event, Delta = 0 otherwise).

Two further response variables are defined: **DeltaEv** taking the values 1 for default, 2 for prepayment and 0 for active/censored and **DeltaDef** = 1 for default, 0 otherwise.

In the competing risk models, the variables **Prepay**, **Default** are used to represent the specific events of early redemption and default respectively with (Prepay = 1, indicating prepayment event, 0 otherwise) and (Default = 1, indicating the default event, 0 otherwise)

To be consistent with the empirical literature, for the purposes of computing the Kaplan-Meier estimator, all mortgage defaults (prepayments) are treated as essentially being censored in the month preceding the one in which they defaulted (prepaid).

The next three sections outline the methodology of the various models that have been implemented for each of the identified themes.

8.4 Focus on Time Varying Covariates

Fundamental to this analysis is the modelling of early termination. Reduced form early termination models are attractive because they allow the researcher to account for heterogeneity among borrowers and to determine how loan characteristics and exogenous economic factors contribute to the borrower's behaviour.

8.4.1 Introduction

The most common cause of early termination is by refinancing and property sale. When interest rates fall, it is to the borrower's financial advantage to lower their total mortgage costs by contracting a new mortgage to pay off the original higher rate mortgage. For most borrowers this is a relatively low cost transaction. Effectively an early termination also occurs when a borrower sells their home and is required to pay off their existing mortgage in order to transfer ownership. Fluctuation in interest rates and changes in house prices drive both these events. Accurate early termination modelling tries to capture the paths of these economic variables and appropriately link them to the borrowers' early termination decision by acknowledging the dependence on the path of these economic variables through the use of time-varying covariates.

Modelling early termination is an active area of research with many models, especially those in more recent studies, based upon the Cox (1972) proportional hazard rate model, which links the probability distribution of a duration variable to a set of covariates.

To partially account for the path of interest rates and house prices, the covariates of the model are allowed to vary with time. Cox showed that a proportional hazard rate model could be adapted to include time-varying covariates through 'episode splitting' that is, continuous time is effectively discretised (split) into (monthly) decision periods and partial likelihood estimation used.

The limitations of the method are that it requires a balanced panel, which is not often available with loan-level mortgage data, and the baseline hazard rate function, which determines how early termination varies with loan age, is factored out of the model.

In this section parametric alternatives that take advantage of the natural panel structure of mortgage data are considered. This feature allows the effective incorporation of time-varying covariates and easily deals with tied events and accounts for censoring in the data.

8.4.2 Model

The problem is approached by adapting the Cox proportional hazard model in order to take account of the discrete structure of data concerning the survival times of mortgages.

A useful measure for mortgage valuation is the hazard rate, which is defined as the probability of default / prepayment in a time period, given survival to that period.

Since default / prepayment leads to early termination of the mortgage, it is reasonable to use a hazard rate model where the underlying variable is the time to mortgage termination, τ , measured as the number of months from origination.

Estimation of the hazard rate is based upon the probability distribution of this variable.

Let T_i be the time when the i^{th} mortgage terminates. The instantaneous early termination risk or hazard rate $\lambda_i(t)$ for the i^{th} mortgage at time *t* is defined as

$$\lambda_i(t) = \lim_{\Delta t \to 0} \frac{P(t \le T_i < t + \Delta t \,|T_i \ge t}{\Delta t} \qquad = \frac{f_i(t)}{S_i(t)} = -\frac{d \ln S_i(t)}{dt}$$

Formally the hazard rate $\lambda_i(t)$ is defined as the probability of failure at time *t*, given survival to time *t*, where the survival function $S_i(t)$, is defined as the probability of surviving to a given time, and is equal to one minus the cumulative distribution function $F_i(t)$.

$$S_i(t) = P(\tau \ge t) = 1 - F_i(t)$$

The probability of failure is given by,

$$f_i(t) = \lim_{\Delta t \to 0} \frac{P(t \le \tau < t + \Delta t)}{\Delta t} = -\frac{dS_i(t)}{dt}$$

By integration the following relationships between the hazard rate, the survival function and the probability density function can be derived,

$$S_i(t) = \exp\left[-\int_0^t \lambda_i(u)du \right] \qquad f_i(t) = \lambda_i(t)\exp\left[-\int_0^t \lambda_i(u)du \right]$$

In the Cox model, the hazard rate is a function of the baseline hazard rate: $\lambda_i(0)$, and the loan characteristics, ν.

$$\lambda_i(t,\nu) = \lambda_i(0)\exp(\beta\nu)$$

The baseline hazard rate function describes how termination varies over time. Covariates serve to shift the baseline hazard rate to account for heterogeneity among borrowers. Often proportional hazard rate models are written in log-hazard form to facilitate the interpretation of the coefficients.

$$\log\lambda_i(t) = \alpha + \beta\nu \qquad \alpha = \log\lambda_i(0)$$

One of the features of the Cox model is that, by using partial likelihood estimation, the baseline hazard rate is factored out of the model.

The baseline hazard rate can be estimated non-parametrically after the estimation of the coefficient vector β

By considering any mortgage rather than the ith mortgage, the notation can be simplified so that equivalently, the hazard rate is assumed to take on the Cox proportional hazards form

$$\lambda(t;x) = \lambda_0(t)\exp[x'(t)\beta]$$

where the baseline hazard at time t is $\lambda_0(t)$, and $x'(t)$ is a vector of time dependent explanatory variables for a mortgage and β is a vector of parameters.

The baseline hazard $\lambda_0(t)$ is the instantaneous default risk of a mortgage with $x'(t) =$ 0. If the covariates are deviations from the mean, then $\lambda_0(t)$ could be interpreted as the hazard rate of an 'average' mortgage.

Thus the model claims an underlying hazard function upon which the individual deviations of the explanatory variables from their mean values act multiplicatively on the instantaneous default risk of an average mortgage.

Even if the time to default could be viewed in principle as a continuous variable, data concerning a mortgage's survival or default is available on a discrete time basis, usually monthly. Indeed what we can typically observe is whether a specified mortgage survives or defaults in a given time interval.

That means the model should be adapted in order to treat time as grouped into disjoint intervals [t_1, t_2), [t_2, t_3), …, [t_k, $t_{k+1} = \infty$)

The discrete time counterpart of the model is thus the probability that the i^{th} mortgage defaults in the interval [t_k, t_{k+1}) conditional upon its survival at the beginning of the same interval

$$\lambda_i(t_k; x_i) = P\{T_i \in [t_k, t_{k+1}) \,|T_i \geq t_k\}$$

This expression can be written as the complement of the conditional probability of surviving the interval given that the mortgage loan was current at the beginning and may be rewritten as follows:

$$\lambda_i(t_k; x_i) = P\{T_i \in [t_k, t_{k+1}) \,|T_i \geq t_k\} = 1 - \exp\left(-\int_{t_k}^{t_{k+1}} \lambda_i(u)du\right)$$

$$= 1 - \exp\left(-\int_{t_k}^{t_{k+1}} \exp[x_i'(t_k)\beta]\lambda_0(u)du\right)$$

$$= 1 - \exp[-\exp(x_i'(t_k)\beta + \alpha_{t_k}]$$

by noting that the covariates vector x_i is a function of t_k not the dummy variable *u* for integration since it is assumed that the independent variables are measured for an entire interval and not for every instant of the interval $[t_k, t_{k+1})$ and defining

$$\alpha_{t_k} = \log\left(\int_{t_k}^{t_{k+1}} \lambda_0(u)du\right)$$

Thus when considering the time variable as grouped in discrete-time intervals, the Cox proportional hazards model degenerates to a binary regression with a complementary –log–log link function. This means that in application with the SAS statistical software package the LOGISTIC procedure with the c-log-log link function can be used to perform the model estimation.

8.4.3 Estimation Technique

Now the Likelihood Function for the estimation can be determined as follows.

Let t_{C_i} be the censoring time of the i^{th} mortgage (the time when the ith mortgage ceases being observed) and t_{B_i} the time when the ith mortgage begins being observed. Then we can write the likelihood function associated with the model as

$$l(\alpha,\beta) = \prod_{i=1}^{N}\left\{\left[1-\exp[-\exp(x_i'(t_{C_i-1})\beta+\alpha_{t_{C_i-1}}]\right]^{\delta_i}\prod_{k=t_{B_i}}^{t_{C_i}-1-\delta_i}\exp[-\exp(x_i'(t_k)\beta+\alpha_{t_k})]\right\}$$

Where $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_T]$ with T denoting the time when the study observation period comes to its end and $\delta_i = 1$ if $T_i \le T$ and 0 otherwise.

The first term in the likelihood expression is equal to one except when a mortgage defaults in t_{C_i}. The second term is the probability that the i^{th} mortgage survives at least until t_{C_i} or in the case of mortgage default in t_{C_i} that it survives at least until t_{C_i-1}

The corresponding log-likelihood function can be expressed as

$$L(\alpha,\beta)=\sum_{i=1}^{N}\left[\delta_i \log\left[1-\exp[-\exp(x_i'(t_{C_i-1})\beta+\alpha_{t_{C_i}-1})]\right]-\sum_{k=t_{B_i}}^{t_{C_i}-1-\delta_i}\exp(x_i'(t_k)\beta+\alpha_{t_k})\right]$$

The Cox proportional hazard rate model has several limitations. Firstly, since the baseline hazard rate has been factored out of the model, this method is not useful for mortgage valuation in which it is necessary to forecast an out-of-sample hazard rate function. Secondly, estimation of the Cox model with time varying covariates requires a balanced panel. Loan level mortgage data does not actually conform to this structure. Finally when covariates are allowed to vary with time, the model can no longer be considered a proportional hazard rate model because the covariates for different individuals vary at different rates.

These problems can be addressed by reference to the complementary log-logistic proportional hazard rate model to include time varying covariates. This model was implemented in the SAS software as outlined and the results are given in Chapter 9.

An alternative to early termination models based upon survival analysis is the use of a random utility model that can be approximated with a mixed logit model.

It is this finding that makes random utility models attractive for application in mortgage analysis where choice between discrete alternatives is to be made.

Since a borrower making the early termination decision is attempting to minimise the cost of their mortgage, a random utility model also provides flexibility for introducing time-varying covariates and other elements that are central to accurate representation of borrower decision processes.

8.4.4 Alternative Model

Several studies in the context of discrete choice modelling have shown that duration models can be estimated as discrete choice models, and in particular that the Cox proportional hazard rate model could be written equivalently as a logit model.

The basis for this equivalence is that in a proportional hazard rate model the odds ratio can be written as a linear function of the covariates

$$\log \lambda(t) = \beta v$$

Under this specification, the assumption is that the time to termination, τ, follows a logistic distribution, but with an alternative definition where the hazard rate is defined in terms of a logistic function

$$\log\left[\frac{\lambda(t)}{1-\lambda(t)}\right] = \beta v$$

it is claimed that these two specifications converge as the hazard rate becomes increasingly small. However in the analysis of mortgages the hazard rate is not necessarily small since for example early termination by prepayment is a relatively frequent event.

Implicit in this latest definition is a random utility model, whereas in the prior specification there are distributional assumptions on the duration variable.

To apply the random utility model to the early termination decision, the borrower makes the decision to prepay or to maintain their mortgage schedule based upon which choice yields the highest utility in each month.

Letting the utility associated with early termination in a given month be a linear function of the loan characteristics and economic variables,

$$U_{it} = \beta v_{it} + \varepsilon_{it}$$

Utility is a latent variable that is not observed, but we do observe the choice made by the borrower in that period. Early termination corresponds to a response variable, y, say, with a value equal to one. If early termination is observed, this implies that the latent utility is positive. $y_{it} = 1 \Rightarrow U_{it} > 0$

Under the assumption that the residuals follow an i.i.d. Gumbel or Extreme value distribution, the probability of early termination is given by a logistic function

$$P(y_{it} = 1) = P(\beta v_{it} + \varepsilon_{it} > 0) = P(\varepsilon_{it} > -\beta v_{it}) = \frac{1}{1 + \exp(\beta v_{it})}$$

An alternative is to assume that the residual follows an i.i.d. Normal distribution

$$P(y_{it} = 1) = \Phi(\beta v_{it})$$

but in the two-choice case, there is little difference between the two assumptions.

8.4.5 Estimation Technique

The coefficients of the model are estimated by maximum likelihood. The likelihood function is calculated by aggregating the probability of the observed choice stream for each individual,

$$\log L = \sum_{n=1}^{N} \ln(P(\{y_{nt} : y_{nt} = 1 \ \forall t = 1 \ldots \tau_n\}))$$

Now all periods prior to the termination event have response variable equal to zero, the last element in the choice stream represents either the month of the early termination or the last month in the observation period, leading to an unbalanced panel.

If early termination is observed, this element is equal to one, otherwise it is set to zero and the loan is right censored. The structure of the choice stream implies that the probability of a choice in any given period is equivalent to the hazard rate for that choice; a choice of early termination for example can only be made if the mortgage loan survives to that period.

The probability of observing a response variable equal to one in any period is equivalent to the hazard rate. If the error term of the utility function is assumed to be independent over time, the log likelihood function can simplify to

$$\log L = \sum_{n=1}^{N} \sum_{t=1}^{\tau_n} y_{nt} \log P(y_{nt} = 1)$$

This model was implemented in the SAS software as outlined and the results are given in Chapter 9.

8.5 Focus on Competing Risks of Mortgage Default and Prepayment

The choice made by a borrower in a period is influenced to some extent by the other alternatives that are available to him in that period, also known as competing risks and the utility obtained from each alternative. Now in basic hazard rate models, competing risks are estimated independently with alternative choices being treated a censored events. By contrast in contingent claims models the borrower is assumed to hold a joint option to terminate that involves both Prepayment and Default. For mortgages we normally define a period to be a month of loan life, and a borrower is assumed to make a choice regarding his mortgage payment in each period.

Each of the choices is associated with a latent utility that is assumed to be a linear function of observed and unobserved factors that include both loan- and borrower-specific characteristics as well as economic variables.

These factors can be both time varying, such as current interest rate, or constant over time such as the term of the loan.

8.5.1 Introduction

As is common in mortgage analysis the model factors are assumed to be the same for all choices from the given initial state. The borrower is assumed to select the choice in each period that is associated with the highest utility with the optimal choice in each period independent of future periods. Prepayment for example is chosen when the utility obtained from this choice is greater than the utility obtained from any other choice.

Multinomial logit models are used to model relationships between a polytomous response variable and a set of regressors. The generalized logit model consists of a combination of several binary logits estimated simultaneously.

Both the generalized logit and conditional logit models are used in the analysis of discrete choice data. In a conditional logit model, a choice among alternatives is treated as a function of the characteristics of those alternatives, whereas in a generalized logit model, the choice is a function of the characteristics of the individual making the choice.

In many situations, and mortgage modelling is no exception, a mixed model that includes both the characteristics of the alternatives and the individual is needed for investigating early termination decision choices.

8.5.2 Multinomial Logit Model

Consider a borrower choosing among *m* alternatives in a choice set during the period. Let P_{jk} denote the probability that borrower j chooses alternative k, let X_j represent the characteristics of the borrower j, and let Z_{jk} be the characteristics of the k^{th} alternative for borrower j The explanatory variables, being characteristics of an borrower, are constant over the alternatives. so the probability that borrower j chooses alternative k is

$$P_{jk} = \frac{\exp(\beta_k' X_j)}{\sum_{l=1}^{m} \exp(\beta_l' X_j)}$$

where $\beta_1 \ldots \ldots \beta_m$ are m vectors of unknown regression parameters.

In the conditional logit model, the explanatory variables Z assume different values for each alternative and the impact of a unit of Z is assumed to be constant across alternatives so the probability that borrower j chooses alternative k in this case is

$$P_{jk} = \frac{\exp(\theta' Z_{jk})}{\sum_{l=1}^{m} \exp(\theta' Z_{jl})}$$

and θ is a single vector of regression coefficients. Then for the mixed logit model that includes both characteristics of the borrower and the alternatives, the choice probabilities are

$$P_{jk} = \frac{\exp(\beta_k' X_j + \theta' Z_{jk})}{\sum_{l=1}^{m} \exp(\beta_l' X_j + \theta' Z_{jl})}$$

$\beta_1, \ldots, \beta_{m-1}$ and $\beta_m \equiv 0$ are the alternative-specific coefficients, and θ is the set of global coefficients

8.5.3 Estimation

As before the coefficients of the model are estimated by maximum likelihood. To illustrate the how the log likelihood is obtained consider for our particular circumstances of a choice set of 3, that is, there are 3 response categories.

If we use category 3 as the baseline category, there are two comparisons to this referent category. All other comparisons can be obtained based on these two comparisons.

Let $\pi_{\mathbf{ij}}$ denote the probability that category j is selected for the i^{th} response, then the logit for the two comparisons are: $\log_{\mathbf{e}} \frac{\pi_{\mathbf{i1}}}{\pi_{\mathbf{i3}}} = \mathbf{X}_{\mathbf{i}}'\beta_1, \ \log_{\mathbf{e}} \frac{\pi_{\mathbf{i2}}}{\pi_{\mathbf{i3}}} = \mathbf{X}_{\mathbf{i}}'\beta_2$

so that

$$\pi_{\mathbf{i}1} = \frac{\exp(\mathbf{X}_{\mathbf{i}}'\beta_1)}{1 + \exp(\mathbf{X}_{\mathbf{i}}'\beta_1) + \exp(\mathbf{X}_{\mathbf{i}}'\beta_2)},$$

$$\pi_{\mathbf{i}2} = \frac{\exp(\mathbf{X}_{\mathbf{i}}'\beta_1)}{1 + \exp(\mathbf{X}_{\mathbf{i}}'\beta_1) + \exp(\mathbf{X}_{\mathbf{i}}'\beta_2)},$$

$$\pi_{\mathbf{i}3} = \frac{1}{1 + \exp(\mathbf{X}_{\mathbf{i}}'\beta_1) + \exp(\mathbf{X}_{\mathbf{i}}'\beta_2)}$$

Hence to estimate parameter vectors β_1, β_2 the Likelihood function is derived by aggregating over each borrower the probability of his observed choice stream and so the probability of the choice stream is simply the product of the probabilities of the event observed in each period.

$$P(Y_i=2)=P(Y_{i1}=0, Y_{i2}=1, Y_{i3}=0) = \pi_{i2} = [\pi_{\mathbf{i1}}]^0 \times [\pi_{\mathbf{i2}}]^1 \times [\pi_{\mathbf{i3}}]^0 = \prod_{\mathbf{j}=1}^{3} [\pi_{\mathbf{ij}}]^{\mathbf{Y_{ij}}}$$

so the likelihood $L = P(Y_1,\ldots Y_n) = \prod_{\mathbf{i}=1}^{\mathbf{n}} \prod_{\mathbf{j}=1}^{3} [\pi_{\mathbf{ij}}]^{\mathbf{Y_{ij}}}$ from which

$$\log L = \log_e P(Y_1,\ldots Y_n) = \sum_{\mathbf{i}=1}^{\mathbf{n}} \left(\sum_{\mathbf{j}=1}^{2} (\mathbf{Y_{ij}} \mathbf{X_i}' \beta_{\mathbf{j}}) - \log_{\mathbf{e}} [1 + \sum_{\mathbf{j}=1}^{2} \exp(\mathbf{X_i}' \beta_{\mathbf{j}})] \right)$$

Since the model is based on the logit model, the likelihood function has a closed form solution and can be estimated using PROC LOGISTIC with the glogit link function to implement this model within the SAS software and the results are given in Chapter 9.

8.5.4 Parametric Accelerated Failure Time models

A number of Accelerated Failure Time models were constructed for a general comparison since competing risks analysis with this type of model is basically the same as with Cox PHM with the key point of treating all events as censored except the one we are focussing on. Only limited output is reported in Chapter 9. note that PROC LIFEREG runs such parametric AFT Models in the SAS statistical software.

In addition an on-the-fly Cumulative Incidence investigation as a means toward more qualitative default risk assessment was also considered, details are given in the next section and reported briefly in Chapter 9

8.5.5 Competing Risks Cumulative Incidence

For competing risks data one often wishes to estimate the cumulative incidence probability of a specific cause of failure. Considering the first cause the corresponding cumulative incidence probability is defined as the probability of dying of Cause 1 before time *t*.

The standard approach for estimating this quantity in the presence of covariates has been based on regression modelling of the cause-specific hazards for all causes, and then combining these to estimate the cumulative incidence probability, for example, by Cox regression models.

Even though it is easy to obtain estimates of the cumulative incidence curve by the above procedure, it may be hard to say much about the effect of specific covariates on the cumulative incidence curve and to identify which covariates have time-varying effects on the cumulative incidence function.

Typically, the important covariate effects for the cause-specific hazards will influence the cumulative incidence curve significantly, but it is possible that an effect on the cause-specific hazard for Cause 1 is countered by an adverse effect on the overall survival, and then, even though the covariate may be highly significant for all cause-specific hazards, it may not affect the cumulative incidence curve for Cause 1.

Also, it is possible that in cause-specific modelling some covariates are influencing the cumulative incidence function, but may not be themselves strongly associated with any of the cause-specific hazards and may therefore be overlooked.

In this section survival data in which each mortgage loan can experience only one of several different types of early termination event over the observation period is considered.

When only one of several different types of event can occur, reference to the probability of these events as ‘competing risks’ is made. With respect to a mortgage loan these are simply the events of default and prepayment. As shown here, modelling competing risks survival data can be carried out using a Cox model, a parametric model or models that utilise Cumulative Incidence.

When considering competing risks survival data a typical ‘cause specific’ approach is adopted whereby the cause specific hazard functions are modelled by the proportional hazards Cox regression model.

Using the semi-parametric proportional hazards model entails estimating a separate Cox regression model for each cause of failure and the hazard corresponding to a specific cause is analysed by considering failures by the other competing causes are treated as censored observations.

Now in many cases the interest is not in the parameter estimates, but rather on the probability of observation of a failure from a specified cause for subjects with a specified set of covariate values. In our case: to observe a default loan given a low income, first time buyer, with a high LTV at origination, or to observe a prepaid loan given, a high income, low LTV, householder moving jobs.

These probabilities, the cumulative incidences may be presented in an unadjusted form if comparisons are required against a benchmark or among baseline groups (strata) or they can be presented in an adjusted format for baseline group differences in other covariates.

Now the Kaplan-Meier product limit method is used to estimate unadjusted cumulative incidence probabilities and the Cox proportional hazards regression is used to estimate adjusted ones. The results of both are usually presented in a tabular or graphical form as will be the case here. The primary drawback of this approach is the requirement that the competing risks are assumed to be independent.

8.6 Focus on Unobserved Heterogeneity and Correlation over Time

In this section the survival modelling approach taking into account the competing risks nature of default and prepayment has been extended to allow for unobserved heterogeneity.

Unobserved heterogeneity arises and can also be modelled through frictions and transactions costs, both financial and intangible. Intangible costs comprise such items as the time spent searching for a new home and finding and accessing refinancing opportunities, and these factors are thought to account for much of the heterogeneity among borrowers. Much of the relevant correlation is due to unobserved factors, such as risk preference, family structure and financial awareness that also accounts for a high degree of heterogeneity

8.6.1 Introduction

Here the model used is the discrete time counterpart of the Cox proportional hazard model extended to allow for this unobserved heterogeneity and correlation over time. The model in this section is again built upon the training data set with the classification performance assessed by means of an out of sample validation test and also comparison is made with that of the logit models previously identified.

To achieve the model build the time variable is considered as grouped in discrete time intervals (months) so as we have seen the Cox PHM degenerates into a binary regression with a complementary log-log link function.

Default and prepayment risks are assumed to be conditionally independent --- default and prepayment regressions can be estimated separately by treating the other type of the termination risk as censored.

However, only the default risk is being considered in this section. A separate linked analysis for prepayment would be for a future analysis undertaking.

Now since the use of a limited set of covariates cannot completely explain the variability in observed times to default and mortgage loans within a similar grouping structure may have dependent survival times due to some unobserved covariate information, then the excess in unexplained heterogeneity conveniently summarized in a ‘frailty’ could result in model inadequacy. This may especially be so in trying to account for why loans with shorter default times are frailer than others.

By grouping together mortgages originated from the same UK Region, then the frailty may reflect the common environment or policy effect on survivals of all the mortgages. The shared frailty model so indicated provides a framework for modelling this dependence and the taking into account of unobservable heterogeneity.

For modelling purposes the frailty has an assumed prior distribution which is updated as the default and prepayment information set evolves over time and the specification of the model is completed by assuming that these group-level frailties are independent and identically distributed with a gamma distribution. The frailty model thus is comprised of a proportional hazard model conditioned on this random effect.

Thus we have the basis in the SAS software to use the PROC LOGISTIC with a c-log-log link to perform the estimation of the no-frailty version (discrete-time version of Cox proportional hazard) and in PROC NLMIXED to perform an estimation of the model with a shared gamma frailty. To obtain the gamma frailty model use is made of the function *GAMINV(p, shape)* in the following way.

This function returns a quantile from the gamma distribution given a probability value, p. Now by noting that it is possible to invert the normal density to obtain a uniform density, and we can obtain the gamma density from the uniform density, proceed as

If u ~ N(0,1) then the cumulative density function returned by *probnorm(u)* is distributed U(0,1) and since probnorm(u) is simply a p-value and so p-values are U(0,1). This then allows a few lines of additional code in the procedure to complete the problem in defining the Maximum likelihood. However, in order to ensure that the normal distribution is accurately approximated then a large number of Gaussian quadrature points are needed (*Qmax = 100* say and *Qtol = 1E-6*) and also for better convergence (*Method = HARDY*) can be applied.

Putting this together gives the following results (Note that the computational burden is great over 4 hours for each application).

8.6.2 Frailty Model

As known, the use of a limited set of covariates could not be completely appropriate in explaining the variability in observed times to default. The excess in unexplained heterogeneity can result in an inadequacy of the model in accounting for why mortgage loans with shorter failure times are frailer than others. To address this problem a random effect can be introduced into a survival model framework and *frailty* is the term coined to denote just such a random effect.

Thus a frailty model essentially entails a proportional hazards model conditioned on the random effect. If the unobserved heterogeneity is assumed to take a multiplicative form then the frailty version of the continuous hazard function can be expressed as

$$\lambda_i(t; x_i \,|V_i = v_i) = v_i \lambda_0(t) \exp[x_i'(t)\beta]$$

Where V_i is a random variable that is assumed to be independent of $x_i(t)$. It varies over the population of mortgages and represents the omitted covariates. It is easy to demonstrate that the discrete version of the above becomes:

$$\lambda_i(t_k; x_i \,|V_i = v_i) = 1 - \exp[-\exp(x_i'(t_k)\beta + \alpha_{t_k})v_i]$$

It is desired to verify the statistical significance of covariates that reflect some of the underlying economic conditions under which the observed mortgages exist. The main feature of these variables is that, varying over time but not over the mortgages, they can be viewed as the output of a stochastic process that is external to the mortgages under study. It should be noted that the inclusion of one or more of these variables could really improve the potential of the model as a predictive tool.

The inclusion of these external variables requires a slight adaptation of the c-log-log model and its frailty extension. This is as a result of the observation that the coexistence of time dummies and these external covariates causes a collinearity problem, with the result that the coefficients of the external variables are always statistically not significant.

The presence of these external covariates thus implies the removal of the dummy time variables and their replacement with the intercept.

Assuming that the values of the external variables change over time then the model becomes

$$\lambda_i(t_k; x_i; z_i) = 1 - \exp[-\exp(\alpha + x_i'(t_k)\beta + z_i'(t_k)\gamma)]$$

And with frailty

$$\lambda_i(t_k; x_i; z_i | V_i = v_i) = 1 - \exp[-\exp(\alpha + x_i'(t_k)\beta + z_i'(t_k)\gamma)v_i]$$

Where α is the intercept, $x_i(t_k)$ is the vector of time-dependent variables for the i^{th} mortgage with β its parameter vector and $z_i(t_k)$ is the vector of time-dependent external covariates with γ its vector of parameters.

8.6.3 Estimation Technique

Log-Likelihood Function for the estimation of the frailty Model can be determined by considering.

If we assume that v_i is a Gamma distribution with mean 1 and variance σ^2 then the log likelihood associated with the discrete model version above can be expressed as follows

$$L(\alpha,\beta,\sigma^2)=\sum_{i=1}^{N}\log\left[\left[1+\sigma^2\sum_{k=t_{B_i}}^{t_{C_i}-1-\delta_i}\exp[x_i'(t)\beta+\alpha_t]\right]^{-\sigma^{-2}}-\delta_i\left[1+\sigma^2\sum_{k=t_{B_i}}^{t_{C_i}-\delta_i}\exp[x_i'(t)\beta+\alpha_t]\right]^{-\sigma^{-2}}\right]$$

Where the variance σ^2 is to be estimated together with the parameter vectors α and β

In the following application interest is centred on comparing the classification performances of the model. For this purpose consideration is given to the expected hazard rate, expressed as a function of its variance σ^2 .

Such a relationship starting from the conditional probability of default in the interval $[t_k, t_{k+1})$ can be derived as follows:

$$P\{T_i\in[t_k,t_{k+1})|v_i\}=\lambda_i(t_k|v_i)\prod_{j=1}^{k-1}[1-\lambda_i(t_j|v_i)]$$

$$=\lambda_i(t_k\left|v_i)\prod_{j=1}^{k-1}\right.\exp[-\exp(x_i'(t_j)\beta+\alpha_{t_j})v_i]$$

$$=\lambda_i(t_k|v_i)\exp[-\sum_{j=1}^{k-1}\exp(x_i'(t_j)\beta+\alpha_{t_j})v_i]$$

$$=\exp[-F_i(t_{k-1})v_i]-\exp[-F(t_k)v_i]$$

Where $F_i(t_k)=\sum_{j=1}^{k}\exp[x_i'(t_j)\beta+\alpha_{t_j}]$ with the definition that $F_i(t_0)=0$

The marginal distribution of default time for the ith mortgage is derived by integrating out the random effect ν_i as follows:

$$P\{T_i \in [t_k, t_{k+1})\} = \int \exp[-F_i(t_{k-1})v_i]h(v_i)dv - \int \exp[-F(t_k)v_i]h(v_i)dv$$
$$= \left(\frac{1}{\sigma^2 F(t_{k-1})+1}\right)^{\sigma^{-2}} - \left(\frac{1}{\sigma^2 F(t_k)+1}\right)^{\sigma^{-2}}$$

where $h(\cdot)$ is the gamma density function.

From which the hazard rate is obtained as: $\lambda_i(t_k; x_i) = 1 - \left(\frac{\sigma^2 F_i(t_{k-1})+1}{\sigma^2 F_i(t_k)+1}\right)^{\sigma^{-2}}$

The estimation of the Gamma frailty model can be performed with the PROC NLMIXED procedure within the SAS statistical software package.

Chapter 9

"*The results of the whole ... [analysis] ... consists in the sum total of the results of all partial ... [analyses] ... ; But these results of separate ... [analyses] ... are settled by different considerations*"

Adapted from Von Clausewitz, 'On Battle'. Part 1, Chapter 9, 'On War' translated by J. J. Graham.

Data Analysis: Findings and Discussion

9.1 Introduction

This Chapter serves to introduce and discuss the data set kindly provided by a local financial institution, as a preliminary to the results of the data analyses performed, along with relevant supplementary data and its source. The data represents a pool of mortgage contracts issued by this single UK originator, operating in the UK mortgage market.

The chapter can be outlined as follows Section 8.1 provides a description of the supplied multivariate event history data, Section 8.2 lists the variables present and their associated univariate statistics, Section 8.3 encompasses a discussion of the origination and payment trends observed in this data and Section 8.4 concludes with the details of the partitioning of the data into a number of sets for the purposes of model building validation and prediction.

9.2 Multivariate Event Histories: Data Description

The mortgage data supplied for use in this thesis consists of the contractual specifications and the individual payment histories of a set of 29,901 mortgages containing both prime and near/sub- prime debt. These Mortgages were originated over a ten year period between Jan 2000 and Jul 2009 and when constructed as a panel the data set consists of 2,010,233 observations. (individual x time points - monthly). The mortgages remained in the portfolio of the financial institution concerned throughout the observation study period, making the research subject to certain confidentiality agreements, and as such the identity of the data source cannot be disclosed and individual reference ID numbers have been re-coded also to provide anonymity.

All are Repayment type mortgages, level payment fully amortised loans most with 25yr terms. Mortgage history period ends in the 3rd Quarter of 2009. The mortgages were all issued for UK homes and all regions of the UK were represented with Scotland being the dominant region.

Quite extensive information was available regarding the active mortgages as indicated below. However, information on redeemed mortgages was somewhat more limited.

9.2.1 Application (Static) Data

Items of Application Information include:

a). For the Individual borrowers: Occupation Status (course grouping); Marital Status; Basic Income at Origination; First Time Buyer Status; Buy-to-Let Status; Self-Certified Status.

b). For the Property Asset: Price; Valuation; Postal Region of Origination.

c). For the Mortgage Deal: Source indicator; Loan Use; High Lending Charge indicator; Application Credit Score; Start Date

d). The Contractual Details of the mortgage include: Mortgage Type; Loan Size and Loan Term; along with Contract Rate and LTV Ratio at Origination.

9.2.2 Performance (Dynamic Event) Data

Items of Information relating to the individual Mortgage's Performance History include:

a). Loan Amortisation: Month End indicator; Redemption Balance; Monthly payment; Interest Bearing Balance.

b). Delinquency: Loan Status indicator; Arrears Level indicator; Arrears amount.

c). Foreclosure: Collections/Possessions indicator; Date Property Possessed; Possession Sale Date; Month possessions process closed.

9.2.3 Data Integrity: Missing Data.

Missing data are usually not the focus of any given study, but researchers frequently encounter missing data when conducting empirical research. It is unlikely that researchers will have complete information for all cases and or items in their studies and in this thesis the situation is the same: some missing data is encountered.

The majority of statistical analyses are not designed for samples with missing data (Allison, 2001), missing data may skew the results, complicate the interpretation of data analyses, and lead to possible loss of information.

How then, will the missing data be handled? If left untreated (i.e., using the default handling of missing data by software), missing data may significantly affect the study results. On the other hand, the application of a selected missing data technique can also intensify the effects of bias, and affect statistical power and the interpretation of the statistical summaries.

By far the standard and most common approach to missing data is to simply omit those cases with missing data on one or more variables used in the analysis and to run the analyses on what remains. This approach is usually called case-wise (or list-wise) deletion, but it is also known as complete case analysis.

Case wise deletion, is the standard treatment of missing data in most statistical packages. The approach assumes that we have the missing completely at random (MCAR) situation a special case of the Missing at Random (MAR) condition. In this case for the sample of reduced size, unbiased coefficient estimates can be obtained without the need for imputation, the coefficient standard errors will also be valid.

Case-wise deletion uses less of the available data as observations that are missing on even a single variable (partially observed records) are dropped. Unfortunately, even when the data are MCAR there is a loss in power using this approach, especially if we have to rule out a large number of subjects.

Alternative approaches identified below may be considered as a replacement for case-wise deletion, though in some of these cases it may be better off to fall back on case-wise deletion.

Mean Imputation represents an approach of *replacing* data that are missing. In mean Imputation each missing value for a given variable is replaced (imputed) by the observed mean of that variable. Mean imputation is well known to produce biased coefficient estimates. Standard errors are also known to be too small, giving narrow confidence intervals.

Mean imputation can be extended to include the inclusion of a dummy variable covariate indicating '*missingness*', here the missingness dummy is used as a diagnostic tool for testing the hypothesis that the missing data are MCAR. If the dummy coefficient is significant then the data are not MCAR. Using dummy variables to code for missing observations was popularized in the behavioural sciences by Cohen and Cohen (1983). However, the approach does not produce unbiased parameter estimates (Jones, 1996), and so is no longer recommended.

Conditional mean imputation is a further refinement. By making use of predicted values from a regression analysis as if they were the actual values obtained. This approach has been around for a long time and has an advantage over mean substitution -at least the imputed value is in some way conditional on other information we have about the case, i.e. on the other covariates - but this still tends to underestimate the standard errors. By substituting a value that is perfectly predictable from other variables, we have not really added more information but we have increased the sample size and reduced the standard error. A further refinement is known as stochastic regression imputation. This approach adds a randomly sampled residual term from the normal (or other) distribution to each the imputed value, this procedure adds a bit of random error to each substitution it does not totally eliminate the problem, but it does reduce it.

Finally, multiple imputations provide a useful strategy for dealing with data sets with missing values. Instead of filling in a single value for each missing value, multiple imputation procedures replace each missing value with a set of plausible values that represent the uncertainty about the right value to impute. These multiply imputed data sets are then analysed by using standard procedures for complete data and the results from these analyses are then combined. The process of combining results from different imputed data sets is essentially an attempt to provide valid statistical inferences that potentially reflect the uncertainty due to missing values.

If we actually observed any given missing data point it would tend to be close to its imputed value, but not exactly equal to it. Hence, imputed values capture only a portion of the variability that would be observed were all the data present.

Perhaps the most useful treatment of modern approaches can be found in Baraldi & Enders (2010), which focuses on newer data imputation methods which replace the missing data with a best guess at what that value would have been if you were able to obtain it.

Of particular note is that to date, little work has been done on the impact of missing data on survival estimates, and none at all on relative survival. (Nur et al 2010).

Allison (2001) suggests that under widely occurring MAR conditions case-wise deletion performs well. Therefore, in our data we adopted case-wise deletion. The following criteria were applied to decide if a case was deleted.

1. If the House Purchase Price, Mortgage Interest Rate, Borrowers Income, Term to Maturity, or a major proportion of the remaining data was missing at loan origination.

2. If there were less than 2 years of actual performance history – these mortgages have not yet had much opportunity to terminate in default or prepayment. In the case of the default event in particular there may exist some delay (from a few months to as long as 2 years) following the borrower's decision to cease payment and the loan being finally classified as in default by the institution.

3. If the loan is indicated to be prepaid in one month but has some other status (e.g. current or delinquent) in the next month – taken to be indicative of a partial prepayment

An additional data cleaning procedure was carried out on the performance dataset due to a missing month in a small proportion of the data. Here the amortisation schedule for each affected mortgage was determined enabling the missing month data to be disentangled from the succeeding month. Note also that in the dataset default and prepayment events correspond to the occurrence of the last paid instalments that are subsequently identified as default or prepayment.

9.2.4 Supplementary Data Sources

The Performance History data is supplemented with downloaded information consisting of Longitudinal Economics Conditions data as provided on-line in spreadsheet form by the Office of National Statistics (ONS). This Time Series information spanning the duration of the observations in the study period, is required to allow projection of incomes, market conditions and mortgage and interest rate movements along with Nationwide Building Society House Price Index information to provide property value trend estimates.

The risk of exogenous ‘trigger’ events is captured by Unemployment Rate and the Divorce Rate. The observation period for these supplemental time series was from Jan 1999 to Jul 2010 in order to allow for any time lag adjustments.

9.3 Empirical Variable Construction

The decisions concerning the choice of variables to be included in the estimations is almost explicitly based on factors of the option-theoretic pricing framework and on borrower characteristics influencing the events of early termination by default and prepayment as previously described. The following two subsections provide a short discussion of these dataset covariates leading to a listing of covariates available for inclusion in the models:

9.3.1 Dependent Variables:

These are simply 0-1 indicator type variables representing the occurrences of Default and Prepayment **Delta**, or the default event only, **DeltaDef**.

In addition to the binary response a trinomial response variable was provided where the coding pattern was Default (1), Prepaid (2), and Active/Censored (3) for **DeltaEv.** These are shown in table 22:

Table 22: Derived Response Variables

Response Variable	**Name**	**Description**
Binary	Delta	=1 if early termination event
		=0 Otherwise
	DeltaDef	=1 if Default Event
		=0 Otherwise
Trinomial	DeltaEv	=1 if Default Event
		=2 if Prepayment Event
		=3 if Active/Censored

9.3.2 Independent or Explanatory Variables:

With the loan level data available the following covariates in Table 23 have potential for inclusion:

Table 23: Potential Explanatory Variables derived from supplied loan data

Explanatory Variables	**Short Name**	**Description**
Equity		
	OLTV	Loan to Value Ratio at Origination
	currentLTV	Current loan to Value Ratio HPI adjusted
	inmnyppay	Likelihood of Prepayment
	NegEquity	Likelihood of Default
	Neq_ind	Negative Equity Indicator =1 if Negative Equity =0 Otherwise
Application		
	loansize	Loan Value £
	purchrsn	Reason for purchase =1 if Buy-to-Let =0 if Owner Occupied

	selfcerts	Self Certification Indicator =1 if self certified =0 Otherwise
	use	Loan Use Indicator =1 if Refinancing =0 Otherwise
	loc	UK Regional Indicator =1 if Scotland =2 if England =3 if Wales
	ftbs	First time Buyer Indicator =1 if first time buyer =0 Otherwise
	rtbs	Right-to-buy Indicator =1 if Right to Buy =0 Otherwise
	src	Source Indicator Borrower type: =1 if Existing =2 if Direct =3 if Introduced
	FRM	Fixed Rate Mortgage Indicator =1 if Fixed Rate =0 Otherwise
	Grp	Group Indicator =1 if Acquired =0 if In-house
	hichgs	High Charges Applied Indicator =1 if higher charges applied =0 Otherwise
	Advnc_score_app	Application Score (In-house)
	Advnc_mar_sta1	Marital Status Indicator =1 if Single =2 if Married =3 if Divorced
	Advnc_occ_cde1	Occupational Code: =1 if Employed =2 if Self Employed =3 if Unemployed
	Inc_ind	Income Indicator =1 if single income =0 if joint income
	mprpo_price	Property Price at Origination
Contract		
	Rate	Mortgage Contract Rate
	mktrate	Market Mortgage Rate
	maxterm	Mortgage loan Term
	Rate_sp	Rate Spread
	Rate_spcu	Rate Spread Cubed
	IR_diff	Interest Rate differential
	Spr_Ten	10 Year Bond Interest Rate Spread
	Spr_Twenty	20 Year Bond Interest Rate Spread
	Vol 10Yr	Volatility
	Vol 20Yr	Volatility

Ability2Pay		
	IncMult	Income Multiple
	DCR	Debt coverage Ratio
MacroEconomic		
	Unemp	Regional Unemployment Rate
	UKDiv_Rate	UK Divorce Rate
	RPI	Retail Price Index
	FTALLSH	Financial Times All Share Stock Index
	BaseRate	UK Bank Base Rate

Temporal		
	duration	
	drn2	Duration squared
	logdrn	Logarithm of duration
	logdrn2	Logarithm of duration squared
	loanage	Mortgage loan age in months
	Quarterly Indicators :	
	Q1	=1 if Mar to May =0 Otherwise
	Q2	=1 if Jun to Aug =0 Otherwise
	Q3	=1 if Sep to Nov =0 Otherwise
	Q4	=1 if Dec to Feb =0 Otherwise
	Surv_Time	Survival time
	app_time	Application time
	exit_time	Termination Event time
Supplementary		
	Anon_bid_nob	Mortgage ID Number
	HPI	Regional House Price Index (Nationwide)
	mktvalprop	Market Value of Property
	Initchgs	Initial Mortgage Charges
	RedempFeePaid	Mortgage Redemption Fee

9.4 Origination and Payment Trends

This section describes the application data and loan histories from the mortgage loan servicer and originator. The characteristics of this dataset are described as a preliminary to the presentation of the results of the data analysis.

In the discussion below the various risk features that have been identified and addressed previously are related to the data observations. Specifically in this section the summary statistics for all loans originated in the observation period of the study are reported in a tabular and graphical form, giving an assessment of the essential features of the dataset before the modelling takes place.

As explained previously the complete dataset as supplied consists of 29 196 individual mortgage loan histories. However to obtain a useable data set for modelling purposes a large number of these mortgages were unusable at the outset 13.9%.

This was further reduced, by and large, due to missing values across important variables (like income, or contract rate), 22.1%, but also to in part to other forms of data entry error (identified only by a close inspection of individual records) and problematic outliers, 15.9%. Whilst this meant that the actual data was reduced substantively - by about a half (48.1%) to 14 041 individual mortgage loans - this still gave a total number of 1 614 715 observations over the period of 115 months for use in model development and testing. The panels of the table below give the variable descriptive statistics and number missing.

Table 24: Variable descriptive statistics and number missing

Variable	Label	N	N Miss	Minimum	Maximum	Mean	Std Dev
currentLTV	currentLTV	29192	4	70.00	100.00	74.20	9.01
NegEquity	NegEquity	29192	0	0.48	0.53	0.50	0.00
LoanSize	LoanSize	29196	0	26500.00	2003280.72	82118.26	74697.70
inmnyPPay	inmnyPPay	29196	0	-1.26	0.61	0.04	0.16
IncMult	IncMult	23386	5810	0.05	90.37	3.46	3.49
DCR	DCR	23386	5810	0.19	356.08	9.16	7.19
rate	rate	29196	0	1.75	8.84	5.24	0.96
mktrate	mktrate	29196	0	1.00	7.74	5.52	0.98
Rate_sp	Rate_sp	29196	0	-3.64	2.90	-0.28	0.89
Rate_spcu	Rate_spcu	29196	0	-48.23	24.39	-0.67	2.81
IR_Diff	IR_Diff	29196	0	0.34	1.35	0.81	0.14
Spr_Ten	Spr_Ten	29196	0	-3.00	0.80	-0.14	0.64
Spr_Twenty	Spr_Twenty	29196	0	-3.91	1.42	0.00	0.81
Vol10Yr	Vol10Yr	29196	0	0.48	0.48	0.48	0.00
Vol20Yr	Vol20Yr	29196	0	0.26	0.26	0.26	0.00
BaseRate	BaseRate	29196	0	0.50	6.00	4.72	0.81
unemp	unemp	29196	0	3.90	7.60	5.76	0.74
UkDiv_rate	UkDiv_rate	29196	0	11.20	14.10	12.99	0.80
RPI	RPI	29196	0	166.60	218.40	185.71	12.78
FTALLSH	FTALLSH	29196	0	1722.28	3454.12	2618.54	452.84

Variable	Label	N	N Miss	Minimum	Maximum	Median	Range
OLTV	OLTV	29192	4	70.00	100.00	85.00	30.00
maxterm	maxterm	29195	1	3.00	1078.00	264.00	1075.00
advnc_score_app	advnc_score_app	29196	0	0.00	603.00	0.00	603.00
duration	duration	29196	0	3.00	115.00	36.00	112.00
Surv_Time	Surv_Time	29196	0	3.00	115.00	36.00	112.00
app_time	app_time	29196	0	1.00	113.00	48.00	112.00
exit_time	exit_time	29196	0	2.00	114.00	35.00	112.00
loanage	loanage	29196	0	5.00	115.00	98.00	110.00

Q1: Mar - May				
Q1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	22384	76.67	22384	76.67
1	6812	23.33	29196	100.00
Q2: Jun-Aug				
Q2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	20934	71.70	20934	71.70
1	8262	28.30	29196	100.00
Q3: Sep-Nov				
Q3	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	21309	72.99	21309	72.99
1	7887	27.01	29196	100.00
Q4: Dec-Feb				
Q4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	22961	78.64	22961	78.64
1	6235	21.36	29196	100.00

Grp				
Grp	Frequency	Percent	Cumulative Frequency	Cumulative Percent
In-house(0)	25225	86.40	25225	86.40
Acquired(1)	3971	13.60	29196	100.00

advnc_mar_sta1				
advnc_mar_sta1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
NMiss	5001	17.13	5001	17.13
single	9584	32.83	14585	49.96
married	10812	37.03	25397	86.99
divorced	3799	13.01	29196	100.00

advnc_occ_cde1				
advnc_occ_cde1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
NMiss	5001	17.13	5001	17.13
Employed(1)	19264	65.98	24265	83.11
Self employed(2)	3992	13.67	28257	96.78
Unemployed(3)	939	3.22	29196	100.00

src				
src	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Existing(1)	3014	10.32	3014	10.32
Direct(2)	7291	24.97	10305	35.30
Introduced(3)	18891	64.70	29196	100.00

loc				
loc	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Scotland(1)	24170	82.79	24170	82.79
England(2)	4762	16.31	28932	99.10
Wales(3)	264	0.90	29196	100.00

ftbs				
ftbs	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
NMiss	5007	17.15	5007	17.15
0(N)	16792	57.51	21799	74.66
1(Y)	7397	25.34	29196	100.00

rtbs				
rtbs	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
NMiss	5012	17.17	5012	17.17
0(N)	22143	75.84	27155	93.01
1(Y)	2041	6.99	29196	100.00

selfcerts				
selfcerts	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0(N)	28022	95.98	28022	95.98
1(Y)	1174	4.02	29196	100.00

hichgs				
hichgs	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0(N)	19605	67.15	19605	67.15
1(Y)	9591	32.85	29196	100.00

inc_ind				
inc_ind	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Joint(0)	10286	35.23	10286	35.23
Single(1)	18910	64.77	29196	100.00

luse				
luse	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
House Purchase (0)	16513	56.56	16513	56.56
Re-Mortgage(1)	12683	43.44	29196	100.00

purchrsn				
purchrsn	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Owner Occupier(0)	3482	11.93	3482	11.93
Buy To Let(1)	25714	88.07	29196	100.00

9.4.1 Loan Origination and Other Payment Trends

The following two tables display the number of loans originated for each of the three regions of the UK, along with the percentage in each region that either defaulted or were prepaid during the period of observation.

Table 25: Year on Year Geographic Distribution of Originations

Year by Year Geographic Pattern of Originations				
Year	**England**	**Scotland**	**Wales**	**# Mortgages**
2000	26	1632	1	1659
2001	20	1240	1	1261
2002	119	1267	2	1388
2003	46	1447	0	1493
2004	356	1299	10	1665
2005	582	1017	34	1633
2006	828	1212	74	2114
2007	119	1418	4	1541
2008	215	989	6	1210
2009	4	73	0	77
Total	2315	11594	132	14041

Table 26: Geographic Distribution of Default and Prepayment

Panel A: Default

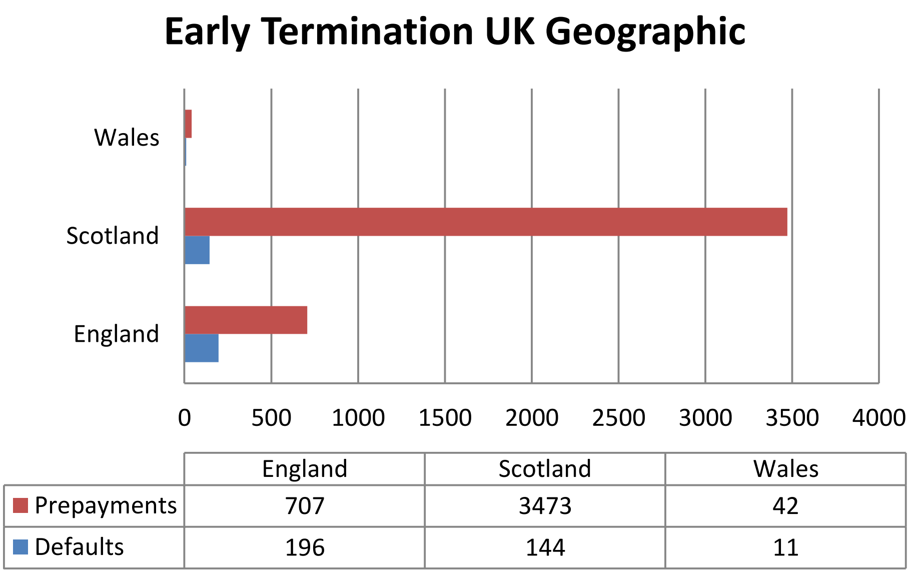
# Mortgages	Defaults	Prepayments
England	196	707
Scotland	144	3473
Wales	11	42
Total	**351**	**4222**

Panel B: Prepayment

% Mortgages	Defaults	Prepayments
England	1.4	5.04
Scotland	1.03	24.73
Wales	0.08	0.3
Total	**2.51**	**30.07**

The mortgages in the study were all issued for UK homes. The mortgage portfolio appears highly diversified geographically. By a consideration of the tables above and the figure below the geographic distribution of early termination shows, as to be expected, a strong concentration in Scotland. This is primarily for the reason that the financial institution supplying the data is based in Scotland.

Figure 28: Geographic Pattern of Early Termination



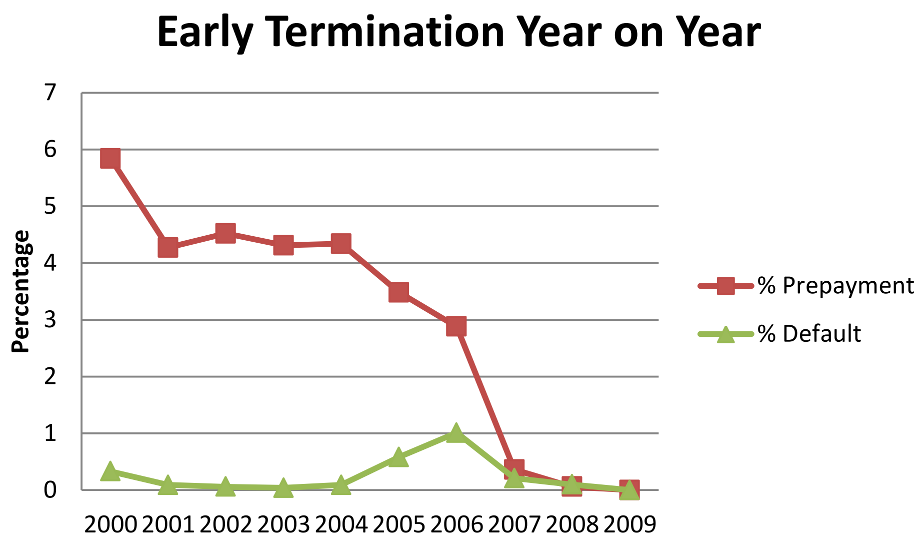
The following table displays the year on year default and prepayment distribution along with the annual total for mortgages originated during the observation period, where it can be seen that the annual total number of originations declines rapidly with the onset of the credit crisis (2007- 2008) with a very dramatic decline in 2009 as mortgage finance appears to have dried up.

Table 27: Year on Year Mortgage Originations

Year by Year Originations: Default and Prepayment Pattern						
Year	**Defaults**		**Prepayments**		**# Mortgages**	**Percent**
	# No.	**Percent**	**# No.**	**Percent**		
2000	46	0.33	820	5.84	1659	11.82%
2001	13	0.09	600	4.27	1261	8.98%
2002	8	0.06	635	4.52	1388	9.89%
2003	5	0.04	605	4.31	1493	10.63%
2004	13	0.09	610	4.34	1665	11.86%
2005	81	0.58	488	3.48	1633	11.63%
2006	142	1.01	404	2.88	2114	15.06%
2007	29	0.21	51	0.36	1541	10.98%
2008	14	0.10	9	0.06	1210	8.62%
2009	0	0	0	0	77	0.55%
Total	**351**	**2.51**	**4222**	**30.06**	**14041**	**100%**

By the end of the observation period under study for this dataset as a whole there were 351 (2.5%) defaulted mortgages, 221 (1.6%) delinquent mortgages, 4222 (30.1%) mortgages were prepaid and 9247 (65.8%) cases remain in and up to date condition.

Figure 29: Year on Year Pattern of Early Termination



In Figure 29 the year on year early termination by prepayment shows a steady decline whereas default – a rare event at the best of times - rises slowly to a peak in 2006.

The average loan size, income multiple and LTI ratio for the dataset as a whole are illustrated in the table of affordability measures below for loans that were either in default, were prepaid or remain current at the end of the observation period. Loans were underwritten according to standard policies in effect over the observation period, including scoring loans using an internally developed mortgage credit scoring model that adds certain borrower and loan characteristics, including LTV, to traditional credit bureau measures, in order to estimate borrower creditworthiness.

Table 28: Affordability features and underwriting terms.

Average Values	Active	Default	Prepaid	All Loans
Loan Size	£88,622.77	£119,124.87	£78,043.38	£94,597.00
Income Multiple	3.5	8	3	4.5
LTI	9	4.6	9	9

Other noteworthy features of the dataset are summarised below in relation to the risk features identified in the previous section. Note that 62% of all loans have LTV ratio below 90% and 28%, roughly a third of borrowers hold loans at origination based on 95% LTV or over.

It is striking that the proportion of Fixed Rate Mortgages to Standard Variable Rate Mortgages issued increased in the years 2004 - 2008 more than doubling in 2008 in the turbulent times before the credit crisis.

Third-Party Broker originations are double those of other channels for mortgages in the period 2004 – 2008 with 4% having low or no documentation of income. These Self-Certified Loans are all in England.

12% loans were for Buy-To-Let investment properties; almost all originated by third party brokers, with the remaining 88% Owner Occupied, half of which are third party broker originations. Approximately a quarter of all loans were to First Time Buyers, all in Scotland.

High lending charges were applied to a third of all loans and 7% of all loans represent Right-to-Buy Mortgages, half of which were for terraced properties around the £50,000 mark.

The median borrower income at mortgage origination was £22,129 with almost half of borrowers in an occupation classified as supervisory/clerical and 65% of borrowers reliant upon a single income source. About one quarter of the active loans and the end of the study period are in negative equity. All the defaulting loans were in negative equity and between 10% and 15% of the prepaid loans had prepayment option values ‘in-the-money’ at point of early termination.

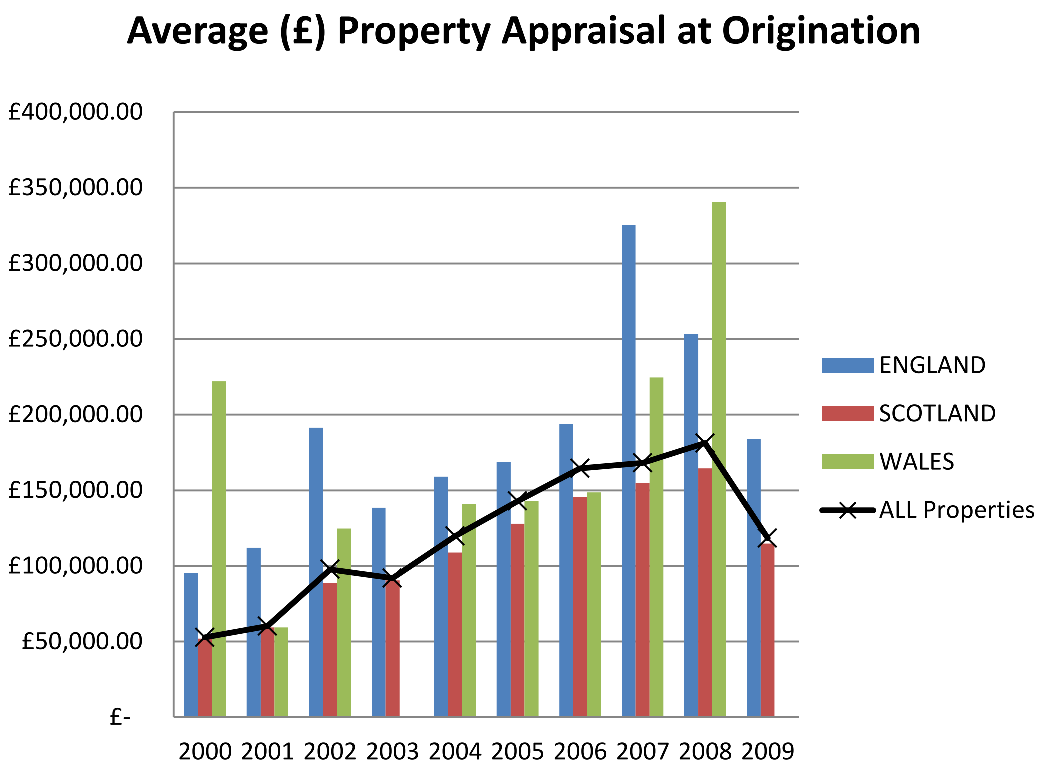
Table 29: Year on Year Average Property Appraisal

Year	Property Appraisal at Origination in £			
	hp_region	Mean		
	ENGLAND	SCOTLAND	WALES	ALL Properties
2000	£ 95,386.04	£ 51,984.38	£ 222,000.00	£ 52,767.06
2001	£ 112,039.40	£ 59,299.33	£ 59,250.00	£ 60,135.77
2002	£ 191,271.69	£ 88,844.24	£ 124,820.50	£ 97,677.69
2003	£ 138,513.57	£ 90,418.26		£ 91,900.10
2004	£ 158,893.67	£ 108,754.48	£ 141,050.00	£ 119,668.90
2005	£ 168,647.28	£ 127,910.87	£ 142,887.50	£ 142,741.12
2006	£ 193,717.48	£ 145,350.58	£ 148,673.84	£ 164,410.99
2007	£ 325,357.87	£ 154,713.11	£ 224,478.50	£ 168,071.83
2008	£ 253,402.62	£ 164,420.10	£ 340,551.67	£ 181,104.42
2009	£ 183,821.50	£ 114,671.68		£ 118,263.88

As can be seen in the table above and accompanying graph below Figure 30., there is a general increase in house price value throughout the period of study, with properties in Scotland remaining entirely below the average and property values in England being consistently higher.

Note that there are few originations in Wales (132) over the period compared to England (2,315) and Scotland (11,594)

Figure 30: Year on Year Average Property Appraisal



Of the loans issued throughout the observation period over 60% were used for house purchase with a peak of 77% in 2007. Approximately 40% of the loans were originated to refinance an existing mortgage loan on the same property.

9.5 Dataset Partitioning

To enable assessments of predictive accuracy the available dataset needs to be partitioned. Partitioning the useable data set yields three mutually exclusive data subsets: a training dataset, a validation dataset and a test dataset. The training dataset is used to train or build/fit the model. Once the model is built on this training data, to get a more realistic estimate of how the model would perform with unseen data, and to find out the accuracy of the model on an unseen dataset, a testing dataset is used.

This is one where the actual values of the target variable (The dependent variable) in this dataset are known. This dataset is known as the validation dataset. The accuracy of the model on the testing data gives a realistic estimate of the performance of the model on '*completely*' unseen data. Thus, another portion of the data which is used neither in training nor in validation needs to be set aside, and this set is known as the hold-out dataset.

Each of these data sets are compiled via a simple random sampling mechanism in which every observation in the main dataset has equal probability of being selected for the partition dataset. The useable dataset used in this study was thus randomly subdivided to create a 50% Training Sample followed by a 35% Validation Sample and finally a 15% Hold-Out Sample keeping the relative proportions of default, prepayment, delinquent and active mortgages the same as indicated in the table:

Table 30: Division of the dataset by sampling

Sampling	50%	35%	15%	100%
	Training Set	Validation Set	HoldOut Set	Full Data Set
Current	4624 (65.85%)	3930 (65.89%)	694 (65.78%)	9248 (65.86%)
Delinquent	111 (1.58%)	93 (1.56%)	17 (1.61%)	221 (1.57%)
Default	176 (2.51%)	148 (2.48%)	27 (2.55%)	351 (2.49%)
Prepaid	2111 (30.06%)	1793 (30.06%)	317 (30.05%)	4222 (30.06%)
Totals	**7022**	**5964**	**1055**	**14041**

9.6 Data Analysis: Initial Exploration

In this section an initial exploratory analysis of the data is described, to determine the significant variables. A discrimination function analysis was performed also, to provide an initial assessment of the quality of the data for the types of investigation to be undertaken and to give a baseline for predictive quality assessments of the other models. In many applications of statistics the modelling strategy can be categorized as follows.

1. *Exploratory / Descriptive* This is a strategy adopted to uncover structure and relationships in data

2. *Confirmatory* This is a strategy for testing whether, and describing how, the 'response' variable (default) varies with a specific 'treatment' variable (LTV), adjusting for another set of variables.

3. *Predictive* This is a strategy used to obtain the best prediction of default from a set of covariates.

In Exploratory/Descriptive analyses there is no specific hypothesis (though one might generate hypotheses). Parsimony is important as is parameter interpretation. The purpose is to examine how the response varies with a set of covariates.

Confirmatory analyses are hypothesis-driven. The predictor of interest and adjustment variables are specified a priori and based on a scientific question. Parsimony and parameter interpretation are important. Such analyses are driven by a scientific question of interest and it is known what you want to adjust for (and why), and adjust for it.

With Predictive analyses the goal is to obtain the best prediction of the response variable from a set of covariates. Here, parsimony - the form of model, and parameter interpretation are not as important as in the other strategies. Forward/backward stepwise procedures with cross-validation are used. The power to predict the response given a new set of covariates is most important. The purpose is to try to minimize out-of-sample prediction error.

In this thesis the primary focus is one of prediction and so the adopted modelling strategy utilises a stepwise variable inclusion methodology (forward/backward inclusion) and the procedure involved is outlined below.

Prior to the model building activity an initial data inspection was performed, included in this the following five assessments were made.

First was to ensure that the provided dataset was as large a sample of cases as possible with sufficient follow-up history to capture mortgage termination events of interest. Second was the deletion of cases with missing observations in the covariates data was performed. Third was to make corrections for trivial errors in the data by imputation (values for a missing month in the dynamic data) and observation value replacement (In the static data for the gender variable, C was replaced with F, due to original data entry error). Fourth was for the complete case data that remained, in that checks for linearity, additivity along with possible variable transformations and plausible interaction terms were performed.

Finally there were checks for outliers or influential observations,(as these may indicate a need for rescaling or truncating highly skewed variables or highlight data errors) and a check of distributional assumptions (stratification violation of proportional hazards assumption, distributional assumption of parametric models).

For the model building exercise the list of potential covariates were divided into meaningful groupings for instance: Equity related; Ability-to-Pay related; Application; Contractual; Macroeconomic and Temporal. Then for a given group of variables, each variable was individually tested and a model was built, stepwise, with the most important variables with a check for any key interactions. Finally the important covariates from the different groups were combined and the process repeated, testing for removal of any the variables, until the final model emerged.

Bearing in mind that in doing the forwards/backwards step-down variable selection, there may be possible loss of information through stepwise techniques as the procedure does not address over fitting. This is the ‘final’ model, next, we validate model and check classification accuracy.

An Internal Validation of the model was undertaken by splitting the data into a training, a validation and a quasi-external validation sample set and making best use of cross-validation methods in the variable assessment stage. Classification Accuracy is to be assessed using the C-statistic and ROC Curve.

After a model has been fitted to a set of observations, adequacy of that model can be investigated using residuals.

Residuals are useful as they are calculated for each individual loan in the dataset and can be plotted to study for apparent patterns indicating lack of fit. Deviance residuals are residuals which should be symmetrically distributed about zero and as such can help identify outlying loans. Deviance residuals can be output from the PHREG procedure in SAS using the RESDEV statement. In this thesis Deviance residuals were plotted against the linear predictor and were randomly scattered and centred symmetrically around a residual value of zero ranging between -3.80 and 3.34 which suggests the data have not been mis-modelled

9.6.1 Data Integrity: Multicollinearity and Variable Selection

Multicollinearity is simply redundancy in the information contained in predictor variables. If the redundancy is moderate, it may make the model coefficients unstable (Allison, 1999) however, it usually only adversely affects the coefficient interpretation (Christensen, 1997), but it has no effect on the model prediction.

Multicollinearity can arise because of improper dummy coding or inclusion of a variable that is directly or indirectly computed from other variables. A high level of multicollinearity raises the standard error of the estimated coefficient, which decreases that coefficients level of significance. In other words, a coefficient may be significant, but because of the presence of other correlated variables its significance may be diminished. Alternatively, a coefficient may be artificially significant and become insignificant when correlated variables are removed.

Multicollinearity can be examined in three ways. By an examination of the simple pairwise (Pearson) correlation between the independent variables:

If an independent variable is highly correlated with other independent variables (r~ 0.85), and is not highly associated with the dependent variable, then that variable was not included analysis.

The second test for multicollinearity can be made using the Variance Inflation Factor (VIF) and Tolerance (VIF is the inverse of the Tolerance (=1-R^2), Griffith and Amerhein, 1997) as an indication of the degree of multicollinearity between the independent variables. VIF represents the amount of inflation in the variance when the multicollinearity of a variable with others exists. Griffith and Amerhein (1997) indicate that a value of VIF that exceeds 10 can lead to a serious multicollinearity. For thesis analyses, VIF values of 10 as a cut-off for multicollinearity was used.

If a variable you think should be statistically significant and is not, then consult the correlation coefficients. The basis of a third test was to examine the absolute correlation between the coefficient estimates. In this thesis a pairwise correlation $r > 0.85$ indicates that one variable has to be excluded from the regression analysis - the least theoretically important of the two. Griffith and Amerhein (1997) suggested that this test reveals a better indication of the linearity between a single variable and the linear combinations of others.

The correlation between coefficient estimates can be outputted by adding an option CORRB to the Proc Logistic SAS statement.

In conclusion, access to the right data is perhaps one of the most challenging exercises facing model development. For instance, the proliferation of stated documentation programs of the late 2000's, wherein incomes were no longer verified, have spill over effects into other variables of interest such as the payment-to-income ratio (PTI).

A simple example illustrates this concern. Once mortgage standards allow borrowers to state their incomes, the likelihood of understating PTI would tend to dampen the resulting statistical relationship as higher income borrowers (lower PTIs) exhibit higher default rates due to the fact that this group's data are tainted by those overstating their incomes. This could effectively wipe out any PTI effect altogether in terms of its statistical importance or at the very least reduce its contribution to explaining default. Ideally then it would be good to obtain data with low error in the explanatory variables. - low measurement error and no missing cases, low/no multicollinearity, low/no influential observations as outliers can affect results significantly. In this thesis standardized residuals were analysed for outliers.

Since logistic regression uses maximum likelihood estimation (MLE) to derive parameters, MLE relies on large-sample asymptotic normality which means that reliability of estimates decline when there are few cases for each observed combination of independent variables. That is, in small samples one may get high standard errors. Goodness of fit measures like model chi-square assume that for cells formed by the categorical independents, all cell frequencies are >=1 and no more than 20% of cells are < 5. In this thesis crosstabs were run to assure this requirement for sampling adequacy was met.

9.6.2 A Discriminant Function/Classification Analysis

Discriminant Analysis was the dominant statistical procedure for default prediction until the 1980's, whereby a multivariate Discriminant Analysis framework was used, pioneered by a method that became popular as the Z-score in credit rating literature.

In effect, Discriminant Analysis is a statistical method that tries to identify a linear combination of various factors which discriminate between "a priori" defined groups - which will include the mortgage loans that are likely to prepay or to default - and the mortgage loans that are not likely to do so. The method of classification relies on maximizing the between-group variance relative to the within-group variance. However, there are obvious problems with the assumptions of this framework, for instance the assumption of normality, the equality of covariance matrices, and no adjustment for any multi-collinearity.

A failure to meet these assumptions may lead to serious misclassifications and consequently give rise to meaningless results. However, Note should be made here that the SAS statistical package used for the analysis does include non-parametric methods for the situation when the data does not have multivariate normal distribution. Since, in this case the validity of this assumption was assessed and was found to be met then it was considered that the standard parametric approach is sufficient for the financial analysis and non parametric methods were not used.

To conduct this initial exploratory data analysis we now simply use the Training Sample - coupled with the chosen predictors and with the dependent variable having been grouped into three main categories of interest Prepaid (2), default (1) still active (0) – to determine that combination of the responses

(Discriminant Function) which best describes each mortgage loan group. Each observation in the dataset is assigned probability of belonging to a given group based upon the distance of its discriminant function from that of each group mean or centroid.

Results

Stage 1: The variables that yield greatest difference between the mortgage loan groups are determined and those that are insignificant are eliminated. Stepwise selection (in SAS) was used in selecting variables for the analysis. In this selection method, there are two probability values that control the variable to be entered (SLENTRY) or removed (SLSTAY) from the model. Using a stepwise method of selection of the 59 possible analysis variables, 35 are to be kept. The significance level of the test for selection and retaining of variables of potential interest was set to be 0.2 and 0.25 so as to be reasonable in obtaining as realistic a number as possible.

A significance level of entry 0.25 has been recommended for stepwise regression analysis (Mickey and Greenland, 1989; Hosmer and Lemeshow, 1989).

A predictor will be considered worthy of inclusion if the test has a p-value of 0.2-0.25 or less. This elimination scheme is being used because all of the predictors in the data set are variables that could be relevant to the model. If the predictor has a p-value greater than 0.25 in a univariate analysis, it is highly unlikely that it will contribute anything to a model which includes other predictors.

For model building consideration will be given to the model which will include all the predictors that had a p-value of less than 0.2-0.25 in the univariate analyses above. In this case it means that we will include the following predictor(s) in our model. Those explanatory variables with reasonable discrimination between groups were found to be as presented in table 31:

Table 31: Explanatory Variables with reasonable discrimination

Explanatory Variables	**Name**	**Description**
Equity Related		
	OLTV	Loan to Value Ratio at Origination
	currentLTV	Current loan to Value Ratio index adjusted
	inmnyppay	Likelihood of Prepayment
	NegEquity	Likelihood of Default
	Neq_ind	Negative Equity Indicator =1 if Negative Equity =0 Otherwise
	loansize	Loan Value £
Intrinsic to Application		
	purchrsn	Reason for purchase =1 if Buy-to-Let =0 if Owner Occupied
	selfcerts	Self Certification Indicator =1 if self certified =0 Otherwise
	loc	UK Regional Indicator =1 if Scotland =2 if England =3 if Wales
	ftbs	First time Buyer Indicator =1 if first time buyer =0 Otherwise
	rtbs	Right-to-buy Indicator =1 if Right to Buy =0 Otherwise
	FRM	Fixed Rate Mortgage Indicator =1 if Fixed Rate =0 Otherwise
	Grp	Group Indicator =1 if Acquired =0 if In-house
	Advnc_score_app	Application Score (In-house)

	Advnc_mar_sta1	Marital Status Indicator =1 if Single =2 if Married =3 if Divorced
Contract Related		
	Rate	Mortgage Contract Rate
	mktrate	Market Mortgage Rate
	maxterm	Mortgage loan Term
	Rate_sp	Rate Spread
	Rate_spcu	Rate Spread Cubed
	IR_diff	Interest Rate differential
	Spr_Ten	10 Year Bond Interest Rate Spread
	Spr_Twenty	20 Year Bond Interest Rate Spread
Ability2Pay Related		
	Inc_ind	Income Indicator =1 if single income =0 if joint income
	IncMult	Income Multiple
	DCR	Debt coverage Ratio
MacroEconomic		
	Unemp	Regional Unemployment Rate
	UKDiv_Rate	UK Divorce Rate
Temporal		
	duration	
	drn2	
	logdrn	
	logdrn2	
	loanage	Mortgage loan age in months
Quarterly Indicators		
	Q1	=1 if Mar to May =0 Otherwise
	Q3	=1 if Sep to Nov =0 Otherwise

Stage 2: Discriminant Analysis (D.A.)

These variables were then checked for reasonable Normality and the dichotomous variables were checked for any low frequency valued entries in cross tabulation with the 3 levels of the Response **depvar**: - DeltaEv; /* 3 Group Response Variable */

Within the analysis the homogeneity of the within-groups covariance matrices were checked before the Discriminant functions were chosen.

As there is no prior knowledge about the actual distribution among the three groups then it is assumed that prior probabilities proportional to the sample sizes are used.

D.A. Summary information: Panel A along with identification, proportion and prior probability of each group Panel B is shown in the table 32.

Table 32: Model Summary information

Panel A

Observations	7019	**DF Total**	7018
Variables	35	**DF Within Classes**	7016
Classes	3	**DF Between Classes**	2

Panel B

Class Level Information					
DeltaEv	**Variable Name**	**Frequency**	**Weight**	**Proportion**	**Prior Probability**
0	_0	4732	4732	0.67417	0.67417
1	_1	176	176	0.025075	0.025075
2	_2	2111	2111	0.300755	0.300755

Table 33: Test of Homogeneity of Within Covariance Matrices

Chi-Square	**DF**	**Pr > ChiSq**
561354.72	1260	<.0001

Since the Chi-Square value is significant at the 0.05 level, the within covariance matrices will be used in the Discriminant function.

Reference: Morrison, D.F. (1976) Multivariate Statistical Methods p252.

The table 34 below gives the classification summary that is the number of observations correctly classified and mis-classified. This classification is done for each of the observations; however, the discrimination function that is used in each case is constructed by taking an observation out of the data.

Thus, every data point is reclassified as if it were a new unknown observation with the result providing a more conservative accuracy assessment.

For these data, it can be seen that all the currently Active (0) observations were correctly classified whilst Default (1) shows an error rate of 24.4% and Prepaid (2) shows an error rate of 11.4%. Overall 4.0% of the observations were mis-classified.

Table 34: Classification Summary with Cross-Validation: Training data

Number of Observations and Percent Classified into DeltaEv				
From DeltaEv	**0**	**1**	**2**	**Total**
0	4732 100	0 0	0 0	4732 100
1	0 0	133 75.57	43 24.43	176 100
2	0 0	240 11.37	1871 88.63	2111 100
Total	4732 67.42	373 5.31	1914 27.27	7019 100
Priors	0.67417	0.02507	0.30076	

Error Count Estimates for DeltaEv				
	0	**1**	**2**	**Total**
Rate	0	0.2443	0.1137	0.0403
Priors	0.6742	0.0251	0.3008	

Now the model has been built and looks reasonable it is subjected to the validation data set and the process is repeated. The results are as expected and largely similar to the Training data - currently Active (0) observations were all correctly classified whilst Default (1) shows an error rate of 27.0% and Prepaid (2) shows an error rate of 11.6%. Overall 4.2% of the observations were mis-classified.

To assess the predictive accuracy the Holdout Sample is now used with the Classification summary shown in the table 35 below:

Table 35: Classification Summary: Holdout data

Number of Observations and Percent Classified into DeltaEv				
From DeltaEv	**0**	**1**	**2**	**Total**
0	711 100	0 0	0 0	711 100
1	0 0	23 85.19	4 14.81	27 100
2	0 0	36 11.36	281 88.64	317 100
Total	711 67.39	59 5.59	285 27.01	1055 100
Priors	0.67417	0.02507	0.30076	

Error Count Estimates for DeltaEv				
	0	**1**	**2**	**Total**
Rate	0	0.1481	0.1136	0.0379
Priors	0.6742	0.0251	0.3008	

According to the output there are 40 completely new observations incorrectly classified giving an overall misclassification error rate of 3.8%

Before moving on to consider a logistic model that compensates for the obvious problems with the assumptions of this framework, previously identified - assumption of normality, the equality of covariance matrices, and no adjustment for any multi-collinearity. Consideration is given to the time varying nature of some of the covariates that were included in this part – in particular those that relate to 'trigger' event processes – unemployment and divorce.

Application information is fixed with respect to time and so poses no problem however the time series information for regional Unemployment and the UK Divorce rate nationwide changes over time. In the analysis here the value of these covariates although time-varying originally were specified as being taken fixed at mortgage origination, in this way 'How informative today's knowledge with respect to eventual default and prepayment of the mortgage loan is explained for in this part'

For a simple assessment the cross validated classification accuracy on the Training sample is compared in table 36 below

Table 36: Model Build Misclassifications: With and without inclusion of Time Varying Covariates (TVC)

Error Count Estimates for DeltaEv				
With TVC	**0**	**1**	**2**	**Total**
Rate	0	0.2443	0.1137	0.0403
Priors	0.6742	0.0251	0.3008	
Without TVC	**0**	**1**	**2**	**Total**
Rate	0	0.2386	0.1217	0.0426
Priors	0.6742	0.0251	0.3008	

While an overall tiny improvement from 4.3% to 4.0% is noted, the misclassification of the individual groups does not show the same consistency during the model build phase. However, for the out – of – sample data misclassifications, consistency is noted.

Table 37: Out-of-Sample Misclassifications: with and without inclusion of Time Varying Covariates (TVC)

Error Count Estimates for DeltaEv				
With TVC	**0**	**1**	**2**	**Total**
Rate	0	0.1481	0.1136	0.0379
Priors	0.6742	0.0251	0.3008	
Without TVC	**0**	**1**	**2**	**Total**
Rate	0	0.1481	0.1167	0.0388
Priors	0.6742	0.0251	0.3008	

To illustrate visually the results a scatter plot of the training data observations is provided and quite clearly shows the improvement that inclusion of time dependent variables in the vector of covariates makes on this occasion.

Figure 31 compared to Figure 32.

Figure 31: Without the Macroeconomic Time Varying Covariate inclusion

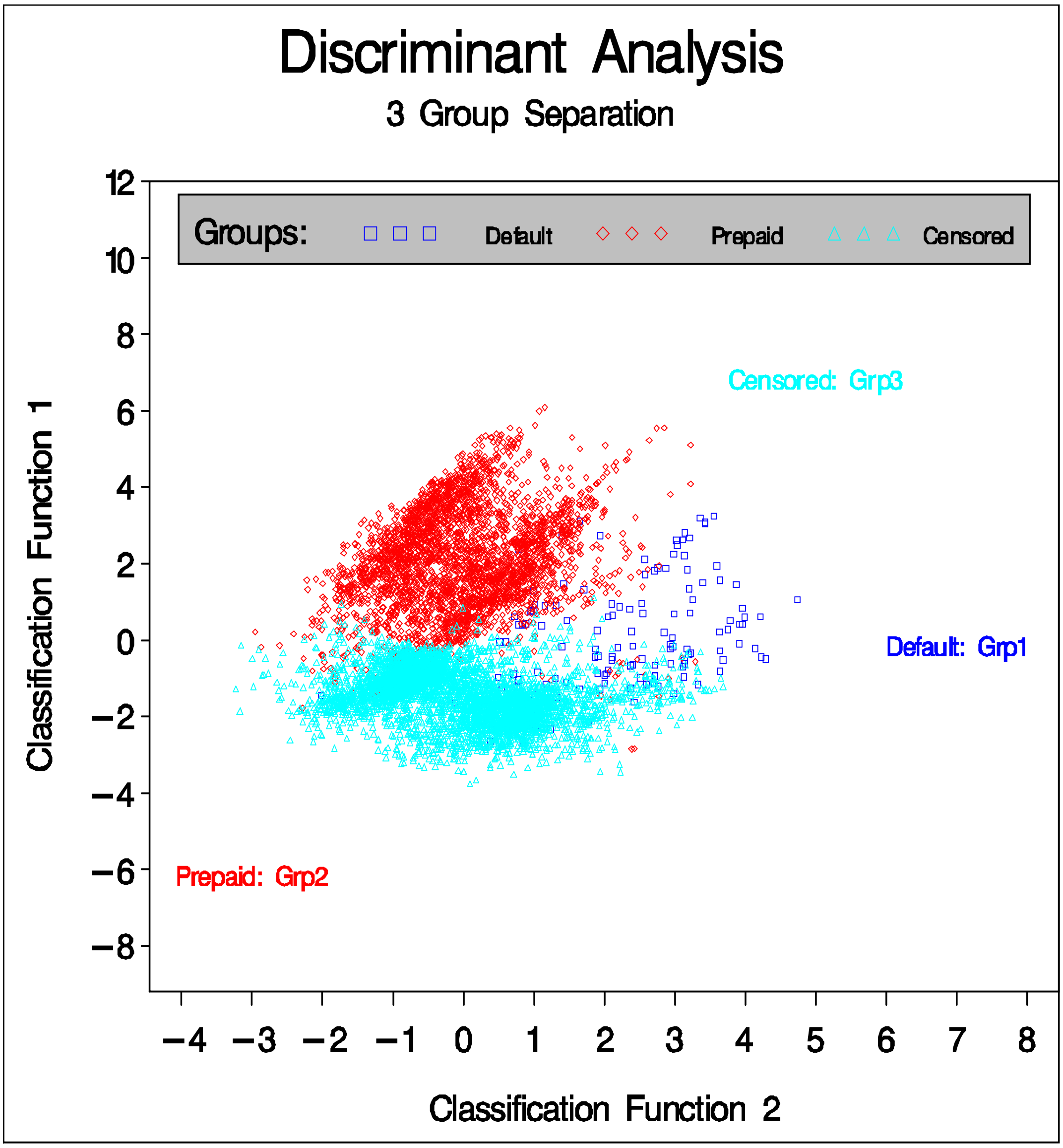
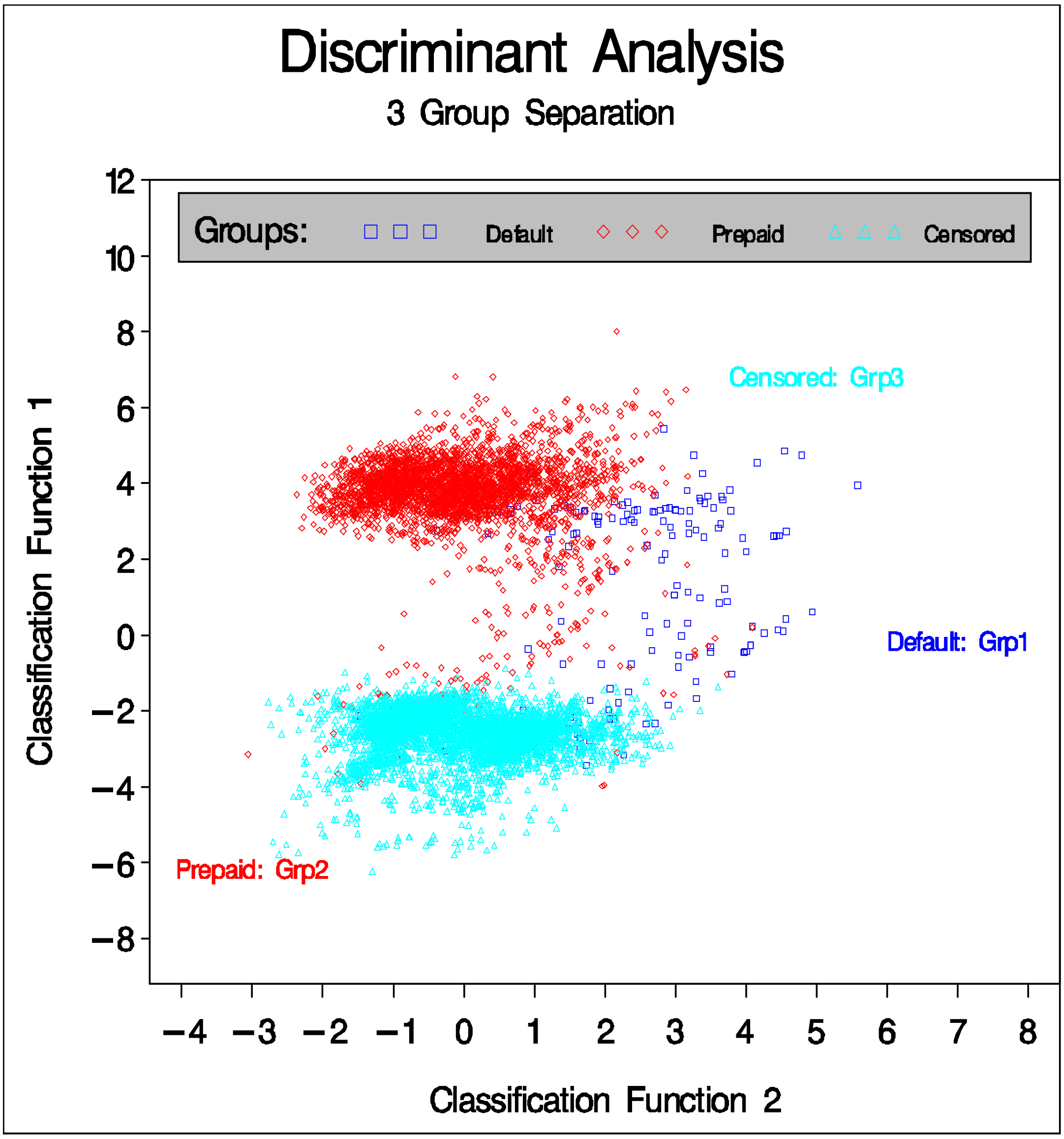


Figure 32: With Macroeconomic Time Varying Covariate inclusion



The next section gives consideration to a logistic modelling basis that compensates for the obvious problems with the assumptions of the discriminant analysis framework, previously identified - assumption of normality, the equality of covariance matrices, and no adjustment for any multi-collinearity.

Post the 1980's discriminant function analysis was increasingly replaced by Logistic Analysis, which is the most commonly used default prediction tool to date.

However, logistic analysis uses the method of maximum likelihood to determine the relationship between binary or ordinal or multinomial response data with a set of explanatory variables. The logistic analysis has two main advantages, like (a) it does not assume normality of financial variables. Instead, it uses a logistical cumulative function to assess the probability of default and prepayment and (b) it also does not assume the equality of covariance matrices.

9.7 Data Analysis Time Varying Covariates

9.7.1 Multiple Logistic Regression Model Results

A multiple logistic regression model was fitted to the mortgage loan data with the previously selected independent predictor variables against two response variable indicators. First a binary response variable (default = 1) with active/censored and prepaid loans coded (default = 0) and secondly in keeping with the competing risks nature of default and prepayment a trinomial response variable, in this case default was coded (delta = 1) with prepaid coded (delta = 2) and active/censored observations (delta = 0) Maximum Likelihood parameter estimates, diagnostic and goodness of fit statistics, and odds ratios were obtained from the fitted model. The fitted model was then used to predict the probability of response by the selected explanatory variables.

Now the application data contains a number of two category type variables, for instance (right-to-buy, first time buyer, income indicator, self-certified and so on) with unbalanced replication of levels across variables the GLM parameterisation option in PROC LOGISTIC CLASS statement was specified – this is needed since SAS by default uses EFFECT parameterisation, and so the GLM coding scheme will estimate the difference in effects of each level compared to the last. As stated previously maximum likelihood method is used to obtain estimates of the model parameters. To enhance the efficiency of this estimation at every step of the fitting process the Newton-Raphson technique was specified. To increase the speed of estimation MODEL statement NOCHECK option was used as it suppresses SAS from checking for infinite parameters. In addition PROC LOGISTIC for each observation computes an event probability, if this probability exceeds a certain 'cut-point' then that observation is predicted to be an event observation otherwise it is predicted as a non-event observation.

Clearly classification accuracy therefore depends upon this critical value. Since we do not know this critical value in advance we can specify a classification table output for a selection of different cut points using for example 0.1 to 1.0 by 0.1 and then select the cut off that gives a reasonable trade-off between sensitivity and specificity subject to the appreciation that increases in sensitivity are at the sacrifice of decreases in specificity. To assist the selection process a ROC curve may be produced to provide a visual demonstration of this trade-off with the area under the curve also calculated in order to quantify the discriminatory power of the model.

In terms of the model build process a number of statistics can be used to subjectively assess model fit. SAS produces two of these:

i) Cox and Snell R^2 – this is a generalised coefficient of determination with the closer the values of R^2 are to 1 the better the fit.
ii) The Nagelkerke R^2, a max rescaled R^2 (that is R^2 / max R^2), the better one to use.

Several other items are also considered useful guides for instance the percent of concordant observations should be close to 100, The C statistic (concordance index equal to the Area under the ROC Curve (AUC)) should be above 0.5 and close to 1, the following fairly arbitrary scale of interpretation will be used here:

Table 38: Area under the ROC curve

AUC (C statistic)	0.5	No Discrimination
	0.6-0.7	Poor Discrimination
	0.7-0.8	Acceptable Discrimination
	0.8-0.9	Very Good Discrimination
	> 0.9	Outstanding Discrimination

Similarly the Gini coefficient (Somers-D in SAS output) should also be above 0.5 and close to 1. Finally there is the Hosmer and Lemeshow goodness of fit test where SMALL values (Pr>ChiSq) are indicative of INADEQUATE fit.

For the Binomial Response:

Default is the event of interest; Prepayment is considered as a censored observation

Default is coded separately from prepayment and active loans.

For the Binomial Logit link function

The Coefficient Estimates and their Standard errors are given in the following tables:

Table 39: Coefficient Estimates: Binary Response Logit Link

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Estimate**		**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-29.3379	***	0.5518	2826.6935	<.0001
IncMult	1	0.1356	***	0.0442	9.4232	0.0021
DCR	1	-0.00252	*	0.00146	2.9597	0.0854
currentLTV	1	0.0715	***	0.00284	635.3041	<.0001
OLTV	1	-0.00998	***	0.00238	17.5588	<.0001
IR_diff	1	0.1841	*	0.0967	3.6215	0.057
inmnyppay	1	-0.5385	***	0.1425	14.2753	0.0002
neq_ind	1	-0.475	***	0.0405	137.7588	<.0001
rate	1	0.3397	***	0.0261	169.5711	<.0001
Q1	1	0.3339	***	0.0294	128.5607	<.0001
rtbs	1	-0.7971	***	0.0486	268.6398	<.0001
maxterm	1	0.00222	***	0.000194	131.6406	<.0001
inc_ind	1	0.4153	***	0.0374	123.1989	<.0001
ftbs	1	0.6222	***	0.0391	252.6765	<.0001
selfcerts	1	0.4393	***	0.065	45.6684	<.0001
advnc_score_app	1	-0.00295	***	0.00032	85.2673	<.0001
FRM	1	-0.3843	***	0.0306	157.7785	<.0001
loc	1	0.2581	***	0.044	34.3981	<.0001
UKDiv_Rate	1	0.8234	***	0.0278	879.2108	<.0001

***** 1% Significance; ** 5% significance; * 10% Significance**

Table 40: Coefficient Estimates: Binary Response c-log-log Link

Analysis of Maximum Likelihood Estimates						
Parameter	**DF**	**Estimate**		**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	1	-29.0766	***	0.5452	2844.4447	<.0001
IncMult	1	-1.139	***	0.1285	78.5884	<.0001
DCR	1	-0.0899	***	0.00878	104.8034	<.0001
currentLTV	1	0.0579	***	0.00549	111.1218	<.0001
OLTV	1	-0.0131	***	0.00464	8.0277	0.0046
IR_diff	1	4.2193	***	0.7452	32.0587	<.0001
inmnyppay	1	2.7524	***	0.48	32.8872	<.0001
neq_ind	1	-0.4569	***	0.081	31.7946	<.0001
rate	1	0.1549	**	0.0737	4.4118	0.0357
Q1	1	0.1087	**	0.0554	3.8516	0.0497
rtbs	1	-0.4647	***	0.0947	24.0966	<.0001
maxterm	1	0.00193	***	0.000337	32.8189	<.0001
inc_ind	1	0.1763	**	0.07	6.3387	0.0118
ftbs	1	0.7582	***	0.0777	95.264	<.0001
selfcerts	1	1.275	***	0.1228	107.8846	<.0001
advnc_score_app	1	-0.0019	***	0.000508	13.9029	0.0002
FRM	1	-0.2671	***	0.0582	21.0431	<.0001
loc	1	0.17	**	0.0792	4.6056	0.0319
unemp	1	-0.4831	***	0.0468	106.7532	<.0001
UKDiv_Rate	1	-0.3063	***	0.0312	96.1656	<.0001

***** 1% Significance; ** 5% significance; * 10% Significance**

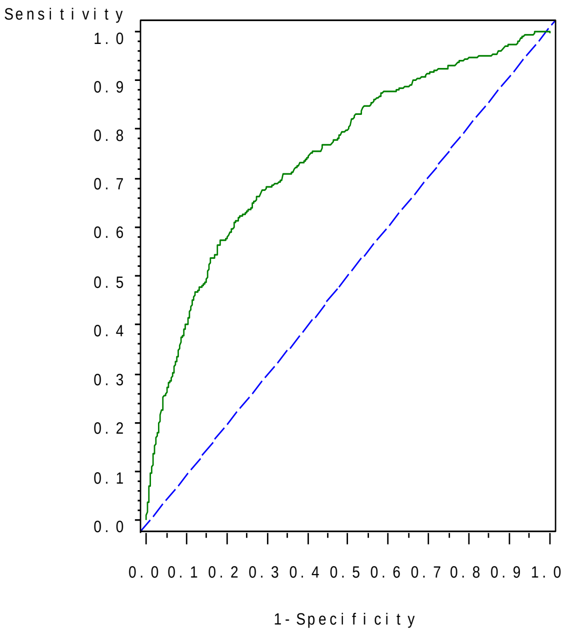
This section reports the results of application of a binary response logistic model both with and without Macroeconomic examples of time varying covariates.

Since we are primarily concerned with prediction and its accuracy only the classification tables and ROC curve output from the analysis are illustrated here and a summarising comment is given at the end.

Table 41: Binary response logit link (with Macro-Economic Factors)

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	68.6	**Somers' D**	0.47
Percent Discordant	21.6	**Gamma**	0.521
Percent Tied	9.8	**Tau-a**	0.005
Pairs	633438000	**c**	0.735

Figure 33: ROC Curve: Binary Response variable, macroeconomic factors and Logit Link



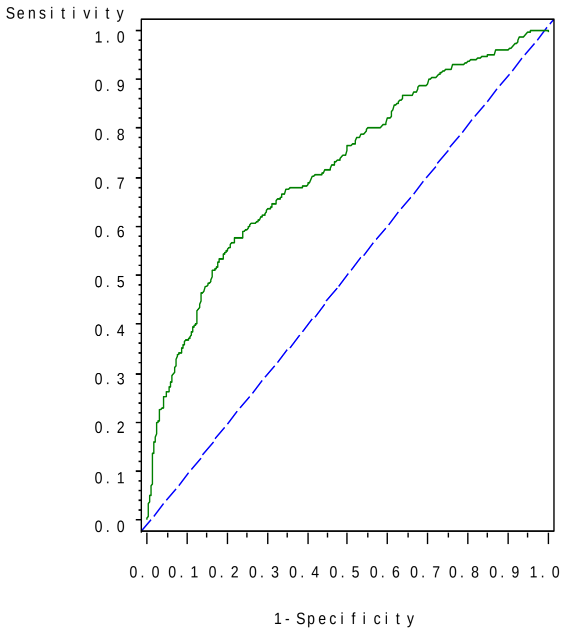
ROC for binary response with Macroeconomic factors and Logit link Area under curve C = 0.735

Discrimination is acceptable as presented with a C statistic of 0.735.

Table 42: Binary response logit link (without Macro-Economic Factors)

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.9	**Somers' D**	0.440
Percent Discordant	26.9	**Gamma**	0.449
Percent Tied	2.2	**Tau-a**	0.026
Pairs	633438000	**c**	0.720

Figure 34: ROC Curve: Binary Response variable, no macroeconomic factors and Logit Link



ROC for binary response without Macroeconomic factors and Logit link Area under curve C = 0.720

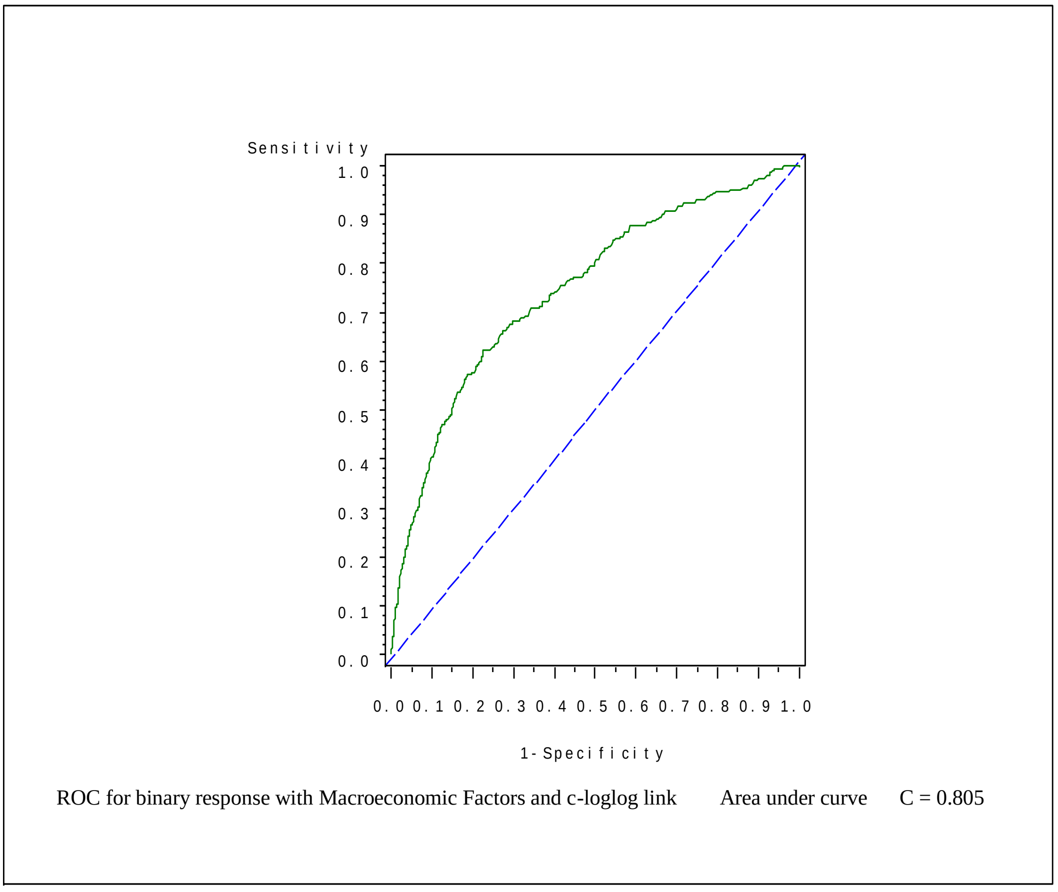
Discrimination is acceptable as presented with a C statistic of 0.720.

Note that there is an improvement to be had by the inclusion of the macro economic variables

Table 43: Binary Response c-log-log link (With Macro-Economic Factors)

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	77.1	**Somers' D**	0.611
Percent Discordant	16	**Gamma**	0.657
Percent Tied	7	**Tau-a**	0.007
Pairs	633438000	**c**	0.805

Figure 35: ROC Curve: Binary Response variable, macroeconomic factors and c-log-log Link

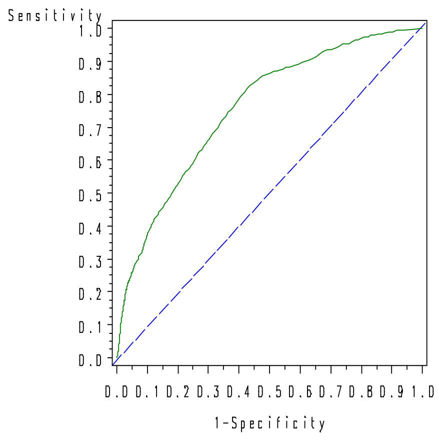


Discrimination is very good as presented with a C statistic of 0.805.

Table 44: Binary response c-log-log link (Without Macro-economic Factors)

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.8	**Somers' D**	0.504
Percent Discordant	20.5	**Gamma**	0.552
Percent Tied	8.7	**Tau-a**	0.006
Pairs	633438000	**c**	0.752

Figure 36: ROC Curve: Binary Response variable, no macroeconomic factors and c-log-log Link



ROC for binary response without Macroeconomic factors and c-log-log link Area under curve C = 0.752

Discrimination is acceptable as presented with a C statistic of 0.752

As can be seen once again the inclusion of macroeconomic variables brings an improvement in the C statistic and so the classification accuracy.

Table 45: C- Statistic Summary

Area under ROC Curve C STATISTIC	**Macroeconomic Variable inclusion**	
Model	**WITHOUT**	**WITH**
LOGIT	0.72	0.735
CLOGLOG	0.752	0.805

Table 45. Demonstrates that overall there is both an improvement when the discrete time Cox PH Model is used and also with the inclusion of the Macroeconomic Variables. (RPI FTALLSH Regional Unemployment and UK Divorce Rate)

9.8 Data Analysis Competing Risks

9.8.1 Results of the Generalised Logit Regression:

Presented here are the results of the Multinomial logit model applied in the context of the competing risks of Default and Prepayment

For the Trinomial response: Competing risks of early termination Default and prepayment coded separately from active (censored) loans.

The Coefficient Estimates and their Standard errors are given in the tables below

Table 46: Coefficient Estimates: Trinomial Response cumulative Logit Link

Analysis of Maximum Likelihood Estimates							
Parameter		**DF**	**Estimate**		**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	**1**	1	3.2754	***	0.0568	3323.5047	<.0001
Spr_Ten		1	-0.2067	*	0.1101	3.5217	0.0606
Loanage		1	-0.0289	***	0.000846	1162.947	<.0001
OLTV		1	-0.00832	***	0.00235	12.5275	0.0004
IncMult		1	-0.534	***	0.0569	88.0628	<.0001
DCR		1	-0.0421	***	0.00327	165.7058	<.0001
Q1		1	0.0817	***	0.0283	8.3065	0.004
currentLTV		1	0.046	***	0.00283	263.1643	<.0001
Spr_Twenty		1	-0.1751	**	0.0865	4.1011	0.0429
IR_diff		1	0.7736	***	0.0972	63.3552	<.0001
logdrn		1	-1.2601	***	0.1088	134.0687	<.0001
logdrn2		1	0.6033	***	0.0558	117.101	<.0001
purchrsn		1	-0.976	***	0.0479	414.6409	<.0001
drn2		1	-0.00015	**	0.000071	4.2237	0.0399
Rate_spcu		1	0.0172	***	0.00668	6.6714	0.0098
Grp		1	1.1709	***	0.0595	386.9466	<.0001
duration		1	-0.0579	***	0.0141	16.8019	<.0001
neq_ind		1	-0.1021	**	0.0416	6.0387	0.014
rate		1	0.0692	**	0.0271	6.5352	0.0106
mktrate		1	0.0614	**	0.0293	4.3767	0.0364
rtbs		1	-0.6286	***	0.0477	173.7692	<.0001
maxterm		1	0.00178	***	0.000191	86.37	<.0001
inc_ind		1	0.3773	***	0.0353	114.3413	<.0001
ftbs		1	0.6518	***	0.038	294.8804	<.0001
selfcerts		1	1.0031	***	0.0648	239.4636	<.0001
advnc_score_app		1	-0.00264	***	0.000312	71.5031	<.0001
FRM		1	-0.352	***	0.0299	138.7487	<.0001
loc		1	0.1522	***	0.0406	14.0651	0.0002
advnc_mar_sta1		1	0.0569	***	0.0154	13.6006	0.0002
unemp		1	-0.7324	***	0.0234	979.3397	<.0001
UKDiv_Rate		1	-0.5204	***	0.015	1195.9906	<.0001

***** 1% Significance; ** 5% significance; * 10% Significance**

Table 47: Coefficient Estimates: Trinomial Response cumulative c-log-log Link

Analysis of Maximum Likelihood Estimates							
Parameter		**DF**	**Estimate**		**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept	**0**	1	6.2232	***	0.1117	3103.5927	<.0001
Intercept	**1**	1	6.832	***	0.1121	3716.429	<.0001
Q3		1	0.0107	*	0.00634	2.8657	0.0905
loanage		1	-0.00485	***	0.00027	322.2527	<.0001
OLTV		1	-0.00991	***	0.00235	17.8162	<.0001
IncMult		1	-0.0351	***	0.00773	20.6588	<.0001
DCR		1	0.00263	*	0.00148	3.1542	0.0757
Q1		1	-0.0657	***	0.00626	110.2337	<.0001
currentLTV		1	-0.011	***	0.000604	330.9374	<.0001
Spr_Twenty		1	0.0484	**	0.0191	6.4019	0.0114
IR_diff		1	0.1978	**	0.0958	4.2657	0.0389
purchrsn		1	0.1481	***	0.0109	183.9366	<.0001
drn2		1	0.000076	***	0.000015	27.6311	<.0001
Rate_spcu		1	-0.00342	**	0.00145	5.5242	0.0188
Grp		1	-0.1752	***	0.0136	165.7374	<.0001
duration		1	-0.0084	***	0.003	7.8151	0.0052
neq_ind		1	0.0566	***	0.00909	38.723	<.0001
inmnyppay		1	0.2679	***	0.0415	41.72	<.0001
rate		1	-0.0455	***	0.00836	29.6468	<.0001
mktrate		1	-0.0891	***	0.0092	93.7884	<.0001
rtbs		1	0.1229	***	0.0101	148.7999	<.0001
maxterm		1	-0.00046	***	0.000043	117.948	<.0001
inc_ind		1	-0.0738	***	0.00704	109.8589	<.0001
ftbs		1	-0.0936	***	0.00777	145.1961	<.0001
selfcerts		1	-0.1229	***	0.0154	64.1108	<.0001
advnc_score_app		1	0.000463	***	0.000054	74.8558	<.0001
FRM		1	0.0829	***	0.00623	176.9821	<.0001
loc		1	-0.071	***	0.00954	55.3658	<.0001
advnc_mar_sta1		1	-0.016	***	0.00311	26.3244	<.0001
UKDiv_Rate		1	-0.1672	***	0.00556	902.6417	<.0001

***** 1% Significance; ** 5% significance; * 10% Significance**

Only the classification summaries are presented here:

Table 48: Trinomial response cumulative logit link

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.3	**Somers' D**	0.55
Percent Discordant	15.3	**Gamma**	0.643
Percent Tied	14.4	**Tau-a**	0.007
Pairs	9092623027	**c**	0.775

Table 49: Trinomial response cumulative c-log-log link

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	74.9	**Somers' D**	0.62
Percent Discordant	12.9	**Gamma**	0.706
Percent Tied	12.1	**Tau-a**	0.008
Pairs	9092623027	**c**	0.810

As can be seen from the two tables (Table 48. and Table 49.) the C value is greatest for the cumulative c-log-log model that is the discrete time version of the Cox PHM provides better classification accuracy with a c=0.810 very good vs c= 0.775 acceptable discrimination.

9.8.2 Parametric AFT Regression

A number of parametric models were estimated and the following tables gives the best fit model ordering based upon the AIC SBC statistic. Both Binomial Response models and Trinomial response models were estimated. As can be seen in both situations the Model using the Gamma link function is most favoured, closely followed in order by the Weibull and LLogistic link function models.

These models were applied in order to determine initial values for the shape and scale parameters required to initialise the version of the c-loglog model with a gamma frailty of the next section.

Table 50: For the binomial response: Model fit based upon AIC SBC

Model	Maximum log-likelihood	AIC	SBC
GAMMA	-37522.41	75116.83	75549.11
WEIBULL	-37529.94	75129.88	75550.16
LLOGISTIC	-37537.03	75144.05	75564.33
LNORMAL	-37744.07	75558.13	75978.41
EXPONENTIAL	-41399.71	82867.43	83275.70
LOG_GAMMA	-67193.35	134458.69	134890.98
TWO_PARAM EXTREME_VALUE	-67207.01	134484.03	134904.30
LOGISTIC	-67219.80	134509.60	134929.88
NORMAL	-67395.93	134861.86	135282.14
ONE_PARAM EXTREME_VALUE	-208487.16	417042.31	417450.58

NB: Gamma distribution chosen with parameter shape 1.8919 and scale 0.1803

For the binomial response situation the better choice appears as the gamma distribution with smallest AIC and SBC values numerically, however subjectively there may be effectively little difference between the three top models except perhaps the number of parameters involved in the models specification and the ease of obtaining them.

Table 51: For the Trinomial response: Model fit based upon AIC SBC

Model	Maximum log-likelihood	AIC	SBC
GAMMA	-36541.79	73155.58	73587.87
WEIBULL	-36548.78	73167.55	73587.83
LLOGISTIC	-36555.46	73180.92	73601.19
LNORMAL	-36746.69	73563.39	73983.66
EXPONENTIAL	-40310.06	80688.12	81096.39
LOG_GAMMA	-65471.5	131014.99	131447.28
TWO_PARAM EXTREME_VALUE	-65484.57	131039.14	131459.42
LOGISTIC	-65496.9	131063.81	131484.08
NORMAL	-65655.89	131381.77	131802.05
ONE_PARAM EXTREME_VALUE	-202629.57	405327.15	405735.42

NB: Gamma distribution chosen with parameter shape 1.9139 and scale 0.1783

For the trinomial response situation we have that the better choice is again the gamma distribution with a similar comment also applying.

In SAS the Maximum likelihood estimators of Parametric Models can be calculated using PROC LIFEREG for a number of survival distribution functions, by specifying the / dist= option on the MODEL statement. The distributions that are supported by SAS 9.1.3 are precisely the ones utilised in this research. The first five models from Exponential, Weibull, Log logistic, Log Normal, Generalised Gamma, Logistic and Normal are also termed Accelerated Failure Time Models, under these models the response variable is modelled as the logarithm of the survival time.

Now when suppressing the log transformation with the NOLOG option for the Exponential or Weibull results in the Extreme value distribution with one and two parameters respectively and lastly if Gamma stipulated with the NOLOG option then the Log Gamma distribution results. This means that 10 different parametric distributions can be generated and fit to the mortgage data to ascertain the best model. To aid the selection the criteria by which they were judged was based upon the value of the Maximum likelihood computed by the PROC LIFEREG and the higher the value of the Maximum likelihood the better is the fit of the model to the data.

However this is not simply the case because of the numbers of parameters involved choosing the best model on this based may in fact lead to weakness in predictability of new cases. It is commonly accepted that the better models are the more parsimonious ones that is adequate explanatory power of the data coupled with the minimum number of parameters. Thus for comparing competing models two other criteria are used to take these factors into account.

These criteria are AIC the Akaike Information criterion and the SBC the Schwarz Bayesian Criterion, they are defined by the formula -2*log-likelihood+ k(#parameters) where k=2 for the AIC and $k = \log(n)$ for n the number of observations for the SBC; the lower the value of these criteria indicating the better models. The selection of the best model to fit the dataset followed these fundamentals and the max likelihood, AIC and SBC determined and then ordered from smallest to largest. The results for the Trinomial response setup - Competing risks and the binary response setup – Early Termination by default give rise to the two tables shown.

By a comparison of both the tables the first four better models fitting the data are GAMMA, WEIBULL, LLOGISTIC and LNORMAL in that order, however just looking at the values of the criteria for numerical superiority, there is very little to distinguish between them.

An additional criterion would be to look at the residuals from the fit of each model to the data and then make a choice, again there are a number of different residuals one can determine and represent visually, in this work the Deviance and Pearson residuals were assessed enabling the decision to choose the Gamma and/or Weibull models.

In actual fact the Gamma Distribution model was chosen but the value of the parameter 'gamma' of the Weibull distribution was also required for the frailty model specification, hence the values of the scale and shape parameters from the PROC LIFEREG were noted and are provided as initial values in the frailty modelling activity for reasons to be discussed.

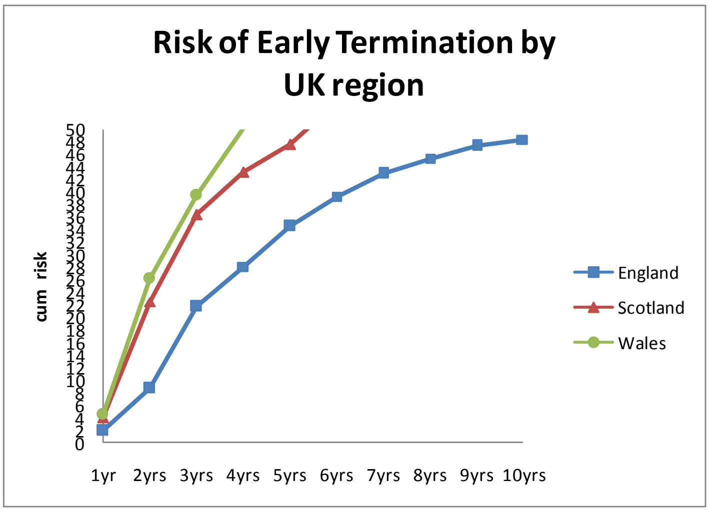
9.8.3 Cumulative Incidence Risk

For the mortgage loans considered in this thesis, this next section presents such an illustrative analysis based upon cumulative incidence.

Table 52: Early Termination risk by Region

duration	cumrisk % loans in England	cumrisk % loans in Scotland	cumrisk % loans in Wales
1yr	1.9	3.9	4.5
2yrs	8.7	22.5	26.1
3yrs	21.7	36.4	39.4
4yrs	28.1	43.1	50.3
5yrs	34.6	47.6	
6yrs	39.3	54.4	
7yrs	43	56.3	
8yrs	45.4	57.3	
9yrs	47.4	65.8	
10yrs	48.4		

Figure 37: Early Termination Risk by UK Region



The cumulative incidence provides risk estimates for the occurrence of a cause specific event in the presence of competing risks.

The highest Cumulative Incidence Probability of 0.484 is reached when t = 10 years when the last observed event occurs. Thus the cumulative risk (marginal probability) for early termination by year 10 is about 0.484 for mortgage originated in England and the cumulative risk (marginal probability) for early termination by year 7 is about 0.563 for mortgages originated in Scotland. When allowing for the competing risks of early termination for Default and Prepayment.

By the end of the second year mortgages in the pool originated in Scotland have more than double the risk of experiencing an early termination event compared to those originated in England similarly those originated in Wales have three times the risk. The cumulative risks of early termination for mortgages originated in Scotland and Wales are typically greater than for those originated in England.

9.9 Unobserved Heterogeneity

In order to obtain unbiased estimates of the classification errors, the predictive ability of the models were checked by means of a hold-out sample test. As previously detailed the aggregate mortgage sample being subdivided randomly into two parts each accounting for 50% of the observations.

The first (Training) set was used to compute the parameter estimates on the basis of the variable selection performed on the entire sample.

The second (Validation) set is used to compute the individual hazards for each of the observed time in years (In the case of an uncensored mortgage from 2000 – 2008).

The yearly time period hazards were then employed to measure the ability of the c-log-log and gamma frailty models to correctly predict the possible outcome in each of the following years (for an uncensored mortgage from 2001 – 2009).

To make a useful comparative analysis with the results provided by a single-period model, estimates for 9 logistic regressions based upon data collected in each of the observed yearly time periods were undertaken. The training and validation sets were then used to assess the classification performance of these logistic regressions. Both for the duration models and the logistic model, the overall classification performance is the mean of the yearly time period results.

As is usual two types of prediction error are defined as follows: A type I error occurs when a mortgage defaults but is predicted to survive; a type II error occurs when a mortgage remains current but is predicted to default. The occurrences of type II misclassifications are counted without considering cases where a mortgage defaults several yearly time periods later than predicted by the model.

As the models are to be used to make a binary prediction, a cut off point for the estimated probabilities of default must be chosen depending upon the relative costs of type I and type II errors.

A concise indicator independent of any cut-off is needed, so for each model an assessment of the area under the receiver operating characteristic (ROC) curve that shows the trade-off between the correct classifications of defaults (sensitivity) and the incorrect classifications of non-defaulted mortgages (1-specificity) made.

This measure has been used in the assessments of the preceding models. But for the future, it may be worthwhile looking at the possible role played by the so called '*lift'* measurement preferred among marketers for evaluating and comparing model performance.

By ranking and tabulating mortgage loans' predicted survival probabilities in ascending order an assessment at each decile can be made. It demonstrates the model's power to beat the random approach or average performance, which is visualized in a *lift gain chart*.

Table 53: Frailty Model comparison of C statistics

Area under the ROC Curves		
Logit	**c-log-log**	**Gamma frailty**
0.726	0.805	0.911

Figure 38: ROC Curve assessment gamma frailty

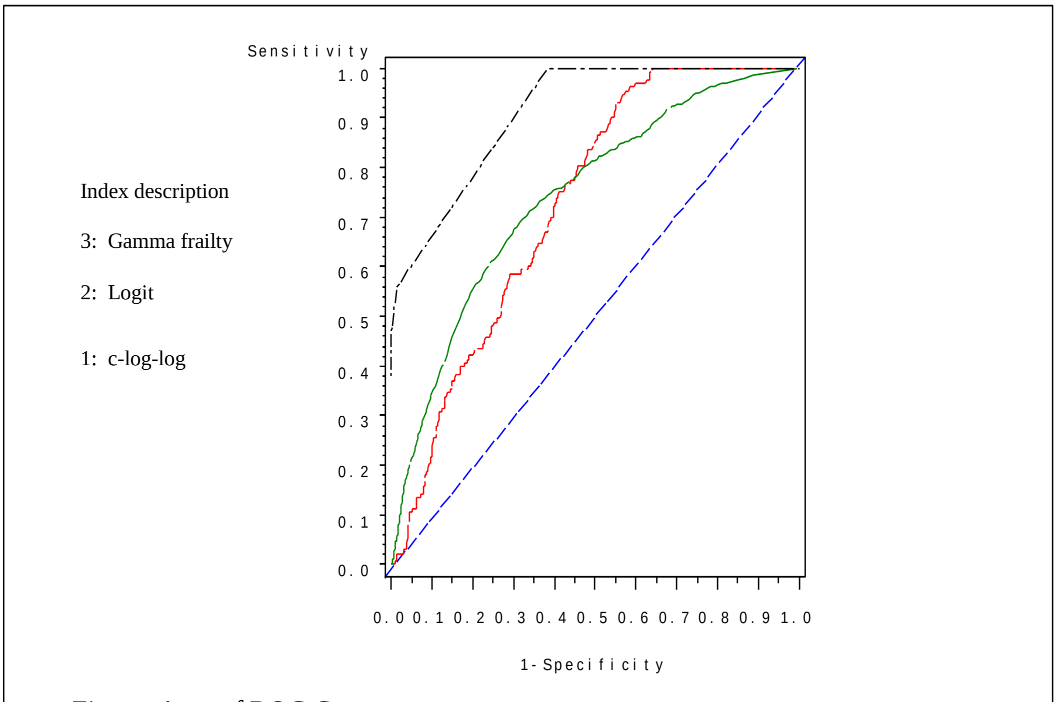


Figure : A set of ROC Curves

As can be seen from the accompanying table (Table 53.) and chart (Figure 38.) the classification accuracy of the Gamma frailty model is almost always better than that of the c-log-log-model.

Chapter 10

‘*This is not the end. It is not even the beginning of the end. But it is perhaps, the end of the beginning*’ ... Winston Churchill at the Mansion House Dinner *09/11/42*

Conclusion

10.1 Summary and Conclusions

The present study has developed and solved a model for the valuation of residential Fixed-Rate Repayment mortgage contracts with specifications that are based upon those underlying such products available to borrowers within the UK mortgage market. Correspondingly the associated insurance related products have also been evaluated.

The thesis opens with reference to the economic significance of the UK mortgage market as a whole with note being made of the profitability of mortgage loan origination alongside the important regulatory issue of setting aside of appropriate capital for the mitigation of residential mortgage credit risk under the Basel II IRB framework. It is important to do this in order to provide constructive background context, and in view of the dissimilarities between UK mortgage products and those available in the US, to commend the worth of developing UK mortgage valuation models and assessing the types of risks involved.

Following on from this, the significant body of literature that has developed particularly in the US but also along different lines in the UK is briefly reviewed with special emphasis placed upon the stochastic economic environment, an option theoretic framework for asset valuation and the behavioural experiences of mortgage borrowers, that shed light upon the validity and stability of structural models.

The discrete time nature of observed mortgage related events and processes informs the type of statistical methodology applied in the production of reduced form models reviewed in the empirical literature. There being a fairly wide degree of overlap and consistency of approach within this research from different application perspectives – for instance taking a comparative view: Commercial versus Residential Mortgages and Prime versus Subprime Mortgages.

In light of the literature review the decisions underpinning the theoretical modelling within the earlier chapters required adoption of a frictionless, competing risks contingent claims valuation model to be taken to provide a final formulation that is retained within the bounds of what is tractable both computationally and mathematically. This means for example that any suboptimal terminal behaviour of borrowers is not considered. In view of these constraints the modelling choices gave rise to a framework based upon the equilibrium model of Cox Ingersoll and Ross (1985a). A two factor stochastic representation of the economy based upon term spot interest rate and house price as state variables was engaged. The valuation function developed was represented in the form of a partial differential equation (PDE) for which there is no closed form solution. This necessitated the use of a numerical technique.

The explicit finite difference was selected since this has been successfully used to solve financial problems involving similar structure and complexity; in addition the algorithm and programming involved are relatively straightforward and seem appropriate to the problems solution.

By suitable variable transformation and discretisation of the state space to ensure stability and convergence of the results, the PDE subject to the application of various boundary, terminal and equilibrium conditions can be solved recursively using a backwards (in time) solution technique. The solution to the problem was enabled through an implementation in Fortran 95 code from within the Salford Plato v2.2 programming and compilation software platform.

Numerical results were generated for a base set of parameters under different economic and contractual specifications, chosen to aid comparison to prior work whilst not being too in consistent with present market realities particularly in terms of the mortgage contract interest rate. A sensitivity analysis was performed and briefly reported. In spite of the rather complex associations existing between the component assets, the evolution of the values of these assets in relation to the characteristics of the economic environment is found to be consistent with economic reasoning.

Finally a note should be given to the issue of failing to consider in detail - particularly methodologically, in light of there being no formal error theory - some relevant features such as the joint termination options or the disincentives - penalties - to early termination in that these may lead to rather questionable conclusions.

Turning now to the supportive empirical investigations undertaken by this work, the literature review informs us that three fundamental avenues of exploration can be discerned and each avenue has been explored by the construction of reduced form models. Each model recognises that a number of loan and non-loan related effects beyond merely the equity in the property and income level of the borrower can influence early termination decisions. Estimations of the parameters of the statistical models developed make use of data supplied by a financial institution issuing mortgage loans consisting of a micro-level sample of individual loan histories over an observation period supplemented by longitudinal census and other economic information. The developed models were estimated using the SAS Statistical Software Platform v9.1.3 Service Pack 3. With default prediction being the primary object of investigation and taking each of the explored avenues individually the findings are:

a). Inclusion of Time-Varying Covariates:

In this section, the Cox (1972) Proportional Hazard Rate Model (one of the commonest reduced-form models in the literature for analysing mortgage performance) was considered, with the benefit that this model brings the ability to associate loan characteristics with default / prepayment rates and its easy adaptation to include time-varying covariates ,(TVC).

In addition the analyses of two fully parametric alternatives that take advantage of the natural panel structure of the supplied mortgage data were also considered. This feature allowed the effective inclusion of time-varying covariates and dealt easily with issues of accounting for tied events and censoring in the data. Inclusion of TVC generally improved the predictive quality of the model as indicated by the use of the ROC curve and C-statistic.

b). The Competing Risks of Early Termination:

In this section a number of models and techniques were used including parametric AFT models and Cumulative Incidence.

The parametric modelling activity was fruitful in that it supplied good initial start values for the parameter estimation involved in the gamma frailty model and was also to provide a qualitative look at default risk measurement from the perspective of a widely used technique borrowed from the field of medicine. Overall the results are consistent with industry consensus and practice.

c). Unobserved Heterogeneity and Correlation Over Time:

As the reviewed literature has shown several empirical estimation studies on mortgage termination risks have been conducted in the framework of survival models. Cox's Proportional Hazards Model (PHM) has been widely used and demonstrated to be effective. In this context the data analysis has been performed to examine whether mortgages are associated in their survival times.

The results obtained from a discrete time adaption of the standard Cox PHM analysis assuming independence of the survival times was representative of the '*No Frailty'* case.

Now, modelling unobserved heterogeneity and correlation over time - *Frailty* - as a representation of an individual mortgage loans' vulnerability to the risk of default or prepayment, captures the total effect of all those influences that impact on the individual mortgage loans' risk of termination that are not included in the baseline hazard function or the information set conveyed by the observed covariates.

Such random effects are not observation specific, but instead are '*shared*' across groups of observations, causing those observations within the same group to be correlated.

The discrete time counterpart of the Cox Proportional Hazards Model (PHM) was extended and used to control for within-group correlation and unobserved heterogeneity It is the results of this proportional hazards model with the unobservable heterogeneity that are representative of the '*Shared-Frailty*' case.

Conforming to most of the existing literature - especially from the medical field - and chosen mainly for its mathematical convenience – as it leads to a closed form for the likelihood function, this model was implemented using the gamma distribution. The empirical results confirm that survival analysis is well placed to portray the dynamics of the mortgage default process. In comparison with single period logit regression this model achieved better performance in terms of predictive accuracy, illustrating that taking into account the unobserved heterogeneity can very favourably improve model predictive accuracy.

The following section provides a discussion of the modelling limitations and provides suggestions for future research.

10.2 Discussion – Limitations Extensions and Future Research

The models introduced in this thesis show great potential for substantially improving the accuracy of valuation models, and increasing our understanding of borrower behaviour. The choice of variables and modelling assumptions attempts to account for as many influences as possible. At the same time, limitations in the amount and quality of available data and the lack of significant 'shocks' to the UK economy in the time period (2000-2006) over which most of the data was collected means there is significant scope for thesis models to be updated and refined over time.

There are many interesting areas of future research activity that can be undertaken to address limitations of and to extend the work completed so far. Possible follow-up activities to overcome weaknesses and extend the structural analytical modelling process might include the following.

1. Undertaking further work using the '*discounting*' approach illustrated in the structural modelling modification chapter. This would be a valuable exercise in linking the economic environment to the structural models available. Unfortunately at the time of completing this work the availability of Leading Economic Indicator data was problematic and only application score information was available in the supplied dataset. If better availability of the economic information and access to behavioural scores was possible, such research could be pursued.

2. The extension of the model to the valuation of mortgage backed securities (MBS) is natural, as the cash flows coming out of a mortgage pool can be directly found with our model. This is especially true for those MBS with rather similar underlying mortgages and even for those with variation on the quality of individual mortgages. Now, the *pass-through MBS* is typically used to distinguish the basic kind of MBS from other type of mortgage backed contracts, this is analogous to the *vanilla option* used to differentiate the basic type of options from the more sophisticated ones.

So, a pass-through security typically has the same type of underlying mortgage loan and these mortgages are similar enough with respect to their maturity and contract interest rate to allow cash flows to be projected as if the pool were a single mortgage.

This is precisely the kind of MBS that the thesis valuation model naturally extends to. To conclude this section a brief outline of how the mortgage valuation of this thesis could be used for the valuation of pass-through MBS is provided.

First note that an estimate of the cash flows coming from a pool of mortgages can be computed by means of the weighted average maturity (WAM) and the weighted average coupon (WAC), which may be used to characterise and describe a pass-through MBS.

This characterisation consists in visualising the pass-through MBS as a single mortgage with maturity given by (WAM) and fixed interest rate given by (WAC). Given this characterization, we can use the approach of this thesis to find the value of a pass-through MBS as if it is a single mortgage. An alternative and also natural approach is to is to find the value of each of the mortgages in the underlying pool by using the WAM, given and the WAC, as the maturity and interest rate respectively of each these mortgages. This approach can be extended further to allow the rating of each of the mortgages with different discount factors, and therefore we may consider a pool of mortgages with different credit quality.

In conclusion, the mortgage valuation model with discounts identified in this thesis may provide a much better and more up-to-date value for each of the individual mortgages in a pool underlying a pass-through MBS.

Unlike models that do not so account for the current conditions of the economy, this model may incorporate conditions like those of the economic crisis of 2007, when the ‘credit crunch’ economy was greatly due to the overpricing of MBS, especially those based on sub-prime mortgages.

3. In following up on the enhanced structural model activities further gain might be achieved by considering other distributional assumptions than Bivariate Normality, to achieve better modelling of the joint interaction of LTV and LTI as a double trigger type mechanism and maybe couple this with a wider applicability - perhaps by application to the countries of the EU on a comparative basis. Loss given default or value at risk may also be useful avenues for exploration.

Possible future research to overcome weaknesses in the data driven modelling process might include the following.

1. Alongside the complete case analysis that was used, a supportive analysis using multiple imputations to investigate the possible influence of variables with larger amounts of missing data and to provide valid inferential alternative results might be undertaken. The MI procedure in SAS can be used to carry out such multiple imputation based on imputations using a Markov Chain Monte Carlo (MCMC) method that assumes multivariate normality to impute missing values. Each of the complete data sets generated could be analysed using PHREG procedure in SAS then the MIANALYZE procedure could be used to generate valid statistical inferences about these parameters by combining results. Multiple imputation as a supportive analysis may then allow more of the mortgage data to be included in the modelling process.

2. It may be argued that predictive models derived using this bespoke dataset are restricted to the mortgage loan population governed by the strict eligibility criteria of the originating financial institution and as such may lack generalisability.

Ideally an external validation - validation on a completely independent dataset is desirable. However where this is not possible an internal validation should be carried out. This is done to assess whether, the model may be transportable and will accurately predict outcome in other mortgage loans not used in the model building process, and thus by inference in future mortgage loans. Further, if external validation is not possible then a possible alternative method instead of data splitting and cross-validation as used in this thesis is an internal model validation based on a statistical re-sampling technique such as a boot strapping method.

3. Borrowers with low credit scores are more likely to default than those with high credit scores; borrowers with small financial buffers are more likely to default than those with big buffers. With the international macroeconomic shocks reflected in recession in the UK economy, the danger that outstanding mortgage debts of marginal homeowners would exceed the true market values of their homes causes considerable concern about the propensity for defaults. Investigation of such variables that measure individual specific information might provide improvement to the model fit. To do so, in our case would require combining the loan-level mortgage data with detailed credit bureau information concerning the borrower's broader balance sheet. This would give a direct way to measure illiquid borrowers: those with high credit card utilization rates.

Results may, by informing central bank monetary policy, partially alleviate household liquidity constraints and lead to a net spending stimulus. By reducing mortgage rates, the Bank of England can increase the net benefits to accessing home equity making it easier for liquidity-constrained households to borrow against their home. The period of low interest rates gives liquidity-constrained homeowners another reason to refinance-they can receive a present value wealth gain by servicing their existing mortgage balance at the lower interest rate.

4. Economic theories regard education and training as investments in human capital that increase the scope of gainful employment and improve net productivity of an individual. The benefit of education and training has been underestimated in most of the studies on default. A possible investigation in an attempt to further fill this gap might be contemplated.

5. One might investigate modelling residential prepayment and default using only the monthly payment status, based upon Markov processes involving state transitions and their associated probabilities akin to the early work of Cyert et al (1962) and latterly by Smith and Lawrence (1995) and Smith, Sanchez and Lawrence (1996).

6. The work of Hsu, Gorfine, and Malone (2007), studies the effect of frailty distribution mis-specification on the marginal regression estimates and hazard functions under assumed gamma distribution. Their results show that the biases are generally low, even when the true frailty distribution deviates substantially from the assumed gamma distribution, suggesting that the gamma frailty model can be a practical choice in real data analyses hence its adoption in the thesis. The standard assumption to use a gamma distribution for the frailty - the random effects- is derived from the fact that the distribution for the frailty term must be positive since the hazard function cannot be negative. However this is a restriction that implies that the dependence is most important for late events.

More generally, the distribution can be stable, inverse Gaussian, or follow a power variance function exponential family. However, other works by Moger and Aalen, (2005) and Wienke et al (2010), use the compound Poisson distribution as the random factor since it allows some individuals to be non-susceptible, which can be useful in the mortgage modelling setting.

Since the gamma distribution was but one choice from three possible plausible distributional forms for the heterogeneity in this thesis, investigating the use of other distributions (Weibull, Log-Logistic and possibly Lognormal) for modelling the unobservable heterogeneity could form the object of a comparative investigation of these distributions.

However, one should nonetheless bear in mind that any approach based on a specific functional form for the distribution of heterogeneity runs the risk of possible sensitivity of its conclusions to the particular functional form used.

7. Future research should continue on identifying variables that can explain the default/prepayment behaviours, suggested by the heterogeneity that exists on the default / prepayment side. As this thesis considered only the total effects of all influences that impact on default prepayment behaviours then perhaps by following the lead of Kau Keenan and Li (2011) a more detailed investigation using a frailty based survival modelling approach utilising stratification to investigate clustering in response might be done. By involving groupings of the supplied data set by LTV, Year of Origination or Regional Location and particularly where the number of events per cluster is small.

10.3 Contributions of the Research to Knowledge

This thesis makes five contributions to knowledge.

1. It has confirmed the theoretical solutions and relationships deduced by recent authors (Kau, J. B., Keenan, D.C., and Kim, T., (1994), Azevedo-Pereira, J.A., Newton, D.P. Paxson, D. A., (2002) and Sharp, N.J., Newton, D.P., and Duck, P.W., (2008)) concerning mortgage valuation and mortgage default. These results were derived using a different solution methodology than previous researchers have followed.

2. It presents an option pricing model of mortgage valuation which unlike previous models assumes house prices follow a mean reverting process rather than a geometric Brownian motion. This shows theoretically that certain elements of local house prices may play a more important role in mortgage valuation than previously shown.

3. It presents a new method to incorporate observable economic conditions into the theoretical options models of mortgage value and default.

4. The thesis presents the first survival model of default hazards for a period including the financial crisis.

5. The thesis presents the first empirical parameterisation of mortgage hazard rates that incorporates unobserved heterogeneity across mortgage borrowers and macroeconomic variables for the UK

10.4 Policy Recommendations

Mortgage loan valuation models as developed in this thesis inclusive of both the credit and interest rate risk is critically important to all market participants; investment bankers, investors, originators, servicers and regulators and academics alike. Investment banking uses these models to design financial innovations - for example credit derivatives, structured products - and hedging strategies.

In addition to the obvious need for accurate valuation models changes in accounting rules provide some of the impetus in enhancing the state of mortgage valuation models. The increase use of risk management techniques and disclosures is another factor. For example, the Basel II Pillar I and Pillar II for capital requirements and the use of Value-at-Risk (VaR) to identify the potential risks of the portfolio result in banks determining the economic value of the mortgage loans. Therefore, banks should find broad applications of such valuation models as the one developed in this thesis in their financial disclosures, profitability and risk management.

Valuation of mortgage loans is also important for non-bank institutions. Investors of whole loans have to evaluate the underlying loan value.

The mortgage insurance business requires consideration of default probabilities in regard to the costs and prepayment likelihoods in regard to the stream of fee income.

For Industry practitioners this thesis valuation model may be useful as a “mark-to-value” proxy for all parties, as expected parameters change (especially interest rate and house price levels, and expected future volatilities), for purposes of determining ‘valued added accounting’, appropriate reserves, and indeed for setting premiums and business decisions.

This thesis provides a model of the valuation of mortgage loans on the balance sheets of banks and other financial institutions. Many banks have originated and held or bought mortgage loans from external sources. They constitute a significant part of their balance sheet. For applied academic researchers the valuation model in this thesis can have a broad range of analytical application, including determining the key rate durations of the mortgages for hedging, and the model can be used to value whole loans and by suitable extension MBS portfolios, mortgage servicing fees, and the cost of default guarantees.

‘*These proceedings are now closed*’
US General Douglas MacArthur, on board *USS Missouri, 02/09/45*

Glossary of Main Terms and Abbreviations

Ability to Pay Theory: Mortgage borrowers tend to avoid defaulting if their current income stream is adequate to fulfil the periodic mortgage payments without too great a burden upon expenditure.

Accelerated Failure Time AFT (Parametric) Regression: A Parametric survival model in which the survival time is assumed to follow a known distribution such as the Weibull or Generalised Gamma Distributions. Data is used to estimate the parameters that fully specify that distribution.

AFEE: Arrangement Fee

AIC and SBC: Choosing the most appropriate parametric model can be difficult. The Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC) are approaches for comparing the fit of models with different underlying distributions by making use of the -2 log likelihood statistic.

American Call Option: A Call option gives the holder the right but not the obligation to buy a particular asset for a given amount, the exercise or strike price, at a specified time in the future, the expiration date. The Option is American if it may be exercised on and at any time prior to the expiration date.

Amortisation Schedule: Tabulation of the changes in interest and principal repayments per month over the length of the mortgage contract. It indicates the gradual elimination of a liability in regular payments in a specified period of time.

Arbitrage: The process of attempting to profit by exploiting the price differences of identical or similar financial instruments

ARM; Adjustable Rate Mortgage

Basel II Banking Supervision: Basel II is a framework to promote the convergence of international capital management and capital standards. The committee on Banking supervision is responsible for setting current capital standards for internationally active banks regarding the amount of regulatory capital that must be put aside to deal with credit risk

BEA: Bureau of Economic Analysis

BHPS: British Household Panel Survey

Censoring: Some Mortgage loans enter the study period after the observation period start and/or leave the study period early, but the majority do not have an early termination event of interest during the entire study period and are termed '*censored*' at the end. The mortgage loans observation time has reached the end of the defined study period and the loan has not yet experienced the early termination event of interest.

CIR: Cox, Ingersoll and Ross, equilibrium model.

CML: Council of Mortgage Lenders

CMO: Collateralised Mortgage Obligations

Contingent Claims: Any set of rights that give the holder the possibility of making future decisions that are capable of affecting the value or timing of cash flows from a venture has value.

CRATE: Mortgage Contract Rate

DCR: Debt Coverage Ratio

Derivative Asset: An asset that is not necessary to describe the underlying real economy

Discriminant Function: Is a mathematical expression, a linear combination of various factors that is used to discriminate between ‘a-priori’ defined groups.

Double Trigger Hypothesis: Attributes default to the joint occurrence of negative equity and a borrower related triggering event that causes an adverse shock to household finances

ETP: Early Termination Penalty

Equity Theory: Borrowers observe a strictly optimising behaviour, maximising financial gain or minimising loss associated with the termination of the mortgage contract.

European Put Option: A Put option gives the holder the right but not the obligation to sell a particular asset for a given amount, the exercise or strike price, at a specified time in the future, the expiration date. The Option is European if it may be exercised only on a specified expiry date.

EAD: Exposure at Default

EL: Expected Loss

Frailty: The representation of an individual mortgage loans’ vulnerability to the risk of default or prepayment, capturing the total effect of all those influences that impact on the individual mortgage loans’ risk of termination that are not included in the baseline hazard function or the information set conveyed by the observed covariates.

FRM: Fixed Rate Mortgage

HPI: House Price Index

In-the-money Option: For a Call option if the Asset price is greater than the Exercise price then the Call option will be in-the-money and will be exercised.

LEI: Leading Economic Indicator

Local Expectations Hypothesis (LEH): The expected change in the value of a pure discount bond is simply equal to the instantaneous interest rate.

LGD: Loss Given Default

LSR: Loan Serviceability Ratio

LTV: Loan to Value Ratio

Mortgage Backed Security (MBS): Mortgage related financial derivative asset. Product is derived from groups of mortgages with common characteristics. Representative bonds are backed upon this pool where the cash flows have been securitised and sold to investors

Mortgage Indemnity Guarantee (MIG): An insurance asset that does not form part of the mortgage package of a borrower. Only the lender benefits from this guarantee. If the borrower chooses to default, the lender will lose the future payments that would have been received. The insurance is used to cover a part of this loss.

NPV: Net Present Value

ONS: Office for National Statistics

Perfect Capital Market: A Market that is devoid of any arbitrage opportunities. In other words, a sufficiently well developed, mature, liquid and costless market.

PD: Probability of Default

PDE: Partial Differential Equation

PHM: Cox Proportional Hazard Model

PTI: Payment to Income Ratio

PV: Present Value

ROC: Return on Capital

***ROC* Curve**: Receiver Operating Characteristic Curve

SAS: Statistical Software Package

Severity: The percentage of the unpaid principal balance that is lost in the event of default

SVR: Standard Variable Rate

Type I Error: This occurs when a Mortgage defaults but is predicted to survive

Type II Error: This occurs when a Mortgage remains current but is predicted to default

TVC: Time Varying Covariates

VLR: Value to Loan Ratio, the inverse LTV

Bibliography

Adams, A., Booth, P, Bowie, D., and Freeth, D., (2003), Investment Mathematics, John Wiley & Sons, Chichester.

Alexander, W. P., Grimshaw, S. D., McQueen, G.R. and Slade, B.A., (2002): Some Loans Are More Equal Than Others: Third-Party Originations and Defaults in the Subprime Mortgage Industry, *Real Estate Economics*, 30(2)., pp 667-697

Allen, L., DeLong, G., and Saunders, A., (2004): "Issues in the Credit Risk modelling of Retail Markets", *Journal of Banking and Finance*, 28(7), pp 727-752.

Allison, P.D., (1982), Discrete-time Methods for the Analysis of Event Histories. In S. Leinhardt (Ed.), *Sociological Methodology*, (pp 61-98) San Francisco.

Allison, P. D. (1995), *Survival analysis using the SAS system: A practical guide,* SAS Institute, Cary, NC

Allison, P.D., (1999): *Logistic Regression Using the SAS System: Theory and Application*. SAS Inst. Inc. Cary, NC.

Allison, P. D. (2001). *Missing Data.* Thousand Oaks, CA: Sage Publications

Ambrose, B.W., Buttimer, R., (2000): "Embedded options in the mortgage contract". *Journal of Real Estate Finance and Economics*, 21(2)., pp 95-111

Ambrose, B.W., Buttimer, R., and Thibodeau, T., (2001): "A new spin on the Jumbo/Conforming loan rate differential". *Journal of Real Estate Finance and Economics*, 23(3), pp 309-335

Ambrose, B.W., Buttimer, R., and Capone, C.A, (1997): "Pricing Mortgage Default and Foreclosure Delay," *Journal of Money, Credit, and Banking*, 29(3), pp 314-325

Ambrose, B.W., Capone, C. A, (1996): "Costs-Benefit Analysis of Single-Family Foreclosure Alternatives," *Journal of Real Estate Finance and Economics,* 13(1), pp105 - 120.

Ambrose, B.W., Capone, C.A, (1998): "Modelling the Conditional Probability of Foreclosure in the Context of Single-family Mortgage Default Resolutions," *Real Estate Economics*, 26(3), pp 391-429.

Ambrose, B. W., Capone, C. A. , (2000): The hazard rates of first and second defaults, *Journal of Real Estate Finance and Economics*, 20(3), pp 275-293.

Ambrose, B. W., Capone, C.A., and Deng, Y., (2000): "The Optimal Put Exercise: An Empirical Examination of Conditions for Mortgage Foreclosure," Forthcoming in *Journal of Real Estate Finance and Economics*, 23(2)., pp 213-234

Ambrose, B. W., LaCour-Little, M., (2001): Prepayment Risk in Adjustable Rate Mortgages Subject to Initial Year Discounts: Some New Evidence, *Real Estate Economics*, 29(2)., pp 305-327

Anderson, R., Vanderhoff, J.,(1999): Mortgage Default Rates and Borrower Race, *Journal of Real Estate Research*, 18(2)., pp 279-289

Archer, W.R., Ling, D.C., (1993): Pricing Mortgage Backed Securities: Integrating Optimal Call and Empirical Models of Prepayment, *Journal of the American Real Estate and Urban Economics Association*, 21(4), pp 373-404.

Archer, W.R., Ling, D.C., (1995): "The Effect of Alternative Mortgage Processes on the Value of Mortgage Backed Securities", *Journal of Housing Research*, 6(2), pp 285-314.

Archer, W. R., Ling, D. C., and McGill, G. A., (1996): The effect of income and collateral constraints on residential mortgage terminations, *Regional Science and Urban Economics*, 26(2), pp 235-261.

Archer, W. R., Elmer, P. J, Harrison, D. M. and Ling, D. C.,(1999): Determinants of Multifamily Mortgage Default, Federal Deposit Insurance corporation, FDIC Working Paper 99-2

Arnold, L., (1992): Stochastic Differential Equations: Theory and Applications, Krieger Publishing Company, Florida.

Aron, J., Muellbauer, J., (2010): *Modelling and Forecasting Mortgage Arrears and Possessions*, Spatial Economic Research Centre, London School of Economics, Paper No SERCDP * , February.

Azevedo-Pereira, J.A., Newton, D.P. and Paxson, D.A.,(2000): "Numerical solution of a two state variable contingent claims mortgage valuation model". *Portuguese Review of Financial Markets* 3(1), pp35-65

Azevedo-Pereira, J.A., Newton, D.P. and Paxson, D. A., (2002): "UK fixed rate repayment mortgage and mortgage indemnity valuation". *Real Estate Economics*, 30(2), pp185-211

Azevedo-Pereira, J.A., Newton, D.P. and Paxson, D.A., (2003): "Fixed-rate endowment mortgage and mortgage indemnity valuation". *Journal of Real Estate Finance and Economics*, 26(2)., pp197-221

Bachelier, Louis., (1900): "Theorie de la speculation", Ann Sci. Ecole Normale Superieure, Vol. 17, pp 21-86. Reprinted in Cootner, P.H. (ed), (1964), The Random Character of Stock Market Prices, MIT Press, Cambridge, Massachusetts.

Bajari, P., Chenghuan, S. C., and Minjung, P.,(2008): An Empirical Model of Subprime Mortgage Default From 2000 to 2007,.NBER Working Paper No. 14625. Available from http://www.nber.org/papers/w14625

Banasik, J., Crook, J.N., and Thomas, L.C.,(1999): Not if but when will borrowers default, *J Opl Res Soc***,** 50, pp 1185-1190

Banasik, J., Crook, J. N. (2005): Credit Scoring augmentation and lean models *J Opl Res Soc.* 56, pp 1072-1081.

Barladi, A. N., Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology,* Vol. 48, pp 5-37

Bellotti T., J. Crook, J. N., (2009): Credit Scoring with macroeconomic variables using survival analysis. *J. Opl Res Soc.* 60, pp 1699-1707.

Bellotti T., J. Crook, J. N., (2012): Loss Given Default Models incorporating macroeconomic variables for credit cards, *International Journal of Forecasting, Special Section 2: Credit Risk Modelling and Forecasting*, 28(1), pp 171–182.

Bennett, P., Peach, R and Peristiani, S.,(2001): Structural Change in the Mortgage Market and the Propensity to Refinance, *Journal of Money, Credit and Banking*, 33(4)., pp955-973

Benninga, S., (2000): *Financial Modelling*, 2nd Ed, MIT Press, Cambridge, Massachusetts.

Berger, A. E., Ciment, M. and Rogers, J. C. W., (1975): Numerical Solution of a Diffusion Consumption Problem with a Free Boundary, *SIAM Journal on Numerical Analysis.* 12(4) pp 646-672

Berk, J., Roll, R., (1988): Adjustable Rate Mortgages: Valuation, *Journal of Real Estate Finance and Economics,* 1(1)., pp 163-184

Bhat, C., (1999): An analysis of evening commute stop-making behaviour using repeated choice observations from a multi-day survey, *Transportation Research: Part B Vol* 33, pp 495–510.

Bhat, C., (2000): Incorporating observed and unobserved heterogeneity in urban work travel mode choice modelling, *Transportation Science* Vol 34, pp 228–238.

Bhutta, N., Dokko, J, and Hui S., (2010): "The Depth of Negative Equity and Mortgage Default Decisions," FEDS Working Paper No. 2010-35. Available at SSRN: http://ssrn.com/abstract=1895493

Bingham, N. H., Kiesel, R., (2006), Risk Neutral Valuation: Pricing and Hedging of Financial Derivatives, Springer Verlag, New York.

BIS, (2001). The Internal Ratings Based Approach. Basel Committee on Banking Supervision, January.

Black, Fisher, Scholes, Myron, (1973): "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, 81(3), pp 637-659.

Bowen, A.,(1994): Housing and the Macroeconomy in the United Kingdom, *Housing Policy Debate*, 5(3)., pp241-251

Boyle, P. P., (1977): Options: A Monte Carlo Approach, *Journal of Financial Economics*, 4(4)., pp323-338.

Boyle, P. P., (1988): " A Lattice Framework for Option Pricing with Two State Variables", *Journal of Financial and Quantitative Analysis*, 23(1), pp 1-12

Brandimarte, P. (2006): Numerical methods in Finance and Economics, John Wiley & Sons Inc. New Jersey

Breedon, F., Joyce, M., (1992): House Prices, Arrears and Possessions, Bank of England Quarterly Bulletin Q2, pp73-179

Brennan, M. J., Schwartz, E. S., (1977): The Valuation of American Put Options. *Journal of Finance.* Vol. 32., pp 449-462.

Brennan, M, J., Schwartz, E.S., (1985): "Determinants of GNMA Mortgage Prices", *Journal of the American Real Estate and Urban Economics Association*, 13(3), pp 209-228.

Brennan, M, J., (1979): The Pricing of Contingent Claims in Discrete Time Models, *The Journal of Finance*, 24(1)., pp 53-68

Breslow, N. E., (1974): Covariance Analysis of Censored Survival Data, *Biometrics,* 30(1), pp89–99.

Brooks, M., Dicks, M. and Pradham, M., (1994): An Empirical model of Mortgage Arrears and Repossessions, *Economic Modelling*, 11(1)., pp134-144

Brueckner, J. K., Follain, J. R., (1988): The Rise and Fall of the ARM: An Econometric Analysis of Mortgage Choice, *The Review of Economics and Statistics,* 70(1), pp93-102.

Brueckner, J.K., (2000): Mortgage Default with Asymmetric Information, *Journal of Real Estate Finance and Economics*, 20(3)., pp 251-274

Brunson, A.L., Kau, J.B., and Keenan, D.C., (2001): A Fixed-Rate Mortgage Valuation Model in Three State Variables. *Journal of Fixed Income,* 11(1), pp 17-27

Bruskin, E., Sanders, A and Sykes, D.:(2001), “The Nonagency Mortgage Market: Background and Overview”, *The Handbook of Mortgage-backed Securities,* edited by Frank J. Fabozzi.

Buist, H., Yang, T.T., (1998),” Pricing the Competing Risks of Mortgage Default and Prepayment in Stochastic Metropolitan Economies”, *Managerial Finance,* 24(1)., pp 110-128.

Burrows, R., (1998): Mortgage Indebtedness in England: an Epidemiology, *Housing Studies*, 13(1)., pp 5-22

Buser, Stephen A., Hendershott, Patric H., (1984): “Pricing Default Free Fixed Rate Mortgages”, *Housing Finance Review*, 3(4), pp 405-429.

Buser, S. A., Hendershott, P. H., and Saunders, A. B., (1985): “Pricing Life of Loan Caps on Default-Free Adjustable-Rate Mortgages, *Journal of the American Real Estate and Urban Economics Association*, 13(3)., pp 248-260

Buser, S. A., Hendershott, P. H., and Saunders, A. B., (1990): “Determinants of the Value of Call Options on Default Free Bonds*”, Journal of Business*, 63(1), pp 33-50.

Calem, P. S., LaCour-Little, M., (2004): Risk-based capital requirements for mortgage loans, *Journal of Banking and Finance*, (28), pp 647–672.

Calhoun, C. A., Y. Deng, (2002): A dynamic analysis of fixed- and adjustable-rate mortgage terminations, *Journal of Real Estate Finance and Economics* 24(2),
pp 9-33.

Campbell, T. S., Dietrich, J. K., (1983),“ The Determinants of Default on Insured Conventional Residential Mortgage Loans”, *The Journal of Finance,* 38(5).,
pp 1569-1582

Campbell, J. Y., Cocco, J. F., (2003): Household risk management and optimal mortgage choice, *The Quarterly Journal of Economics,* Vol. 118, pp 1449-1494.

Capozza, D. R., Thomson, T.A.,(2005): Optimal stopping and losses on subprime mortgages, *Journal of Real Estate Finance and Economics,* 30(1), pp115–131.

Capozza, D. R., Kazarian, D. and Thomson, T.A., (1997): "Mortgage Default in Local Markets," *Real Estate Economics*, 25(4)., pp 631-655

Capozza, D. R., Kazarian, D. and Thomson, T.A., (1998): The Conditional Probability of Mortgage Default, *Real Estate Economics*, 26(3)., pp 359-369

Carron, A.S., Hogan, M., (1988): The Option Valuation Approach to Mortgage Pricing, *Journal of Real Estate Finance and Economics,* 1(1)., pp 131-149

Chambers, M.S., Garriga, C. and Schlagenhauf, D., (2008): Mortgage Innovation, Mortgage Choice, and Housing Decisions, Federal Reserve Bank of St. Louis, Working Paper Nov/Dec 2008.

Chatterjee, A., Edminster, R.O. and Hatfield, G.B.,(1998): An Empirical Investigation of Alternative Contingent Claims Models for Pricing Residential Mortgages. *Journal of Real Estate Finance and Economics*, 17(2), pp 139-162.

Chen, A., Ling, D., (1989): Optimal Mortgage Refinancing with Stochastic Interest Rates., *Journal of the American Real Estate and Urban Economics Association* Vol. 17, pp 278-299.

Chen, R. R., Scott, L., (1992): Pricing Interest Rate Options in a Two Factor Cox Ingersoll-Ross Model of the Term Structure. *Review of Financial Studies,* Vol. 5, pp 613-636

Chen, X., Ender, P., Mitchell, M. and Wells, C. (2003). *Regression with SAS*, Available at http://www.ats.ucla.edu/stat/sas/webbooks/reg/default.htm

Cheng, K., (2003): An Overview of Barrier Options, Global Derivatives Working Paper. Available at: http://www.global-derivatives.com

Childs, P.D., Ott, S.H. and Riddiough, T.J., (2001): “Valuation and information acquisition policy for claims written on noisy real assets”. *Financial Management,* Summer 2001, pp 45-75

Childs, P.D., Ott, S.H. and Riddiough, T.J.,(2002a): “Optimal valuation of noisy real assets”. *Real Estate Economics,* 30(3), pp 385-414

Childs, P.D., Ott, S.H. and Riddiough, T.J., (2002b): “Optimal valuation of claims on noisy real assets: theory and application”. *Real Estate Economics*, 30(3): pp 415-443

Childs, P.D., Ott, S.H. and Riddiough, T.J., (2004): “Effects of noise on optimal exercise decisions: the case of risky debt secured by renewable lease income”. *Journal of real estate finance and economics* 28(2/3): pp109-121

Chinloy, P. (1998): The Probability of Prepayment, *Journal of Real Estate Finance and Economics*, 2(2)., pp267-283

Chinloy, P. (1993): Elective Mortgage Prepayment: Termination and Curtailment, *Journal of the American Real Estate and Urban Economics Association*, 21(3)., pp 313-332

Chinloy, P and Megbolugbe, I. F., (1994): “Real Estate Market Institutions: Implications for the United States”, *Housing Policy Debate*, 5(3), pp 381-399.

Chomsisengphet, S., Pennington-Cross, A.,(2006): “The evolution of the subprime mortgage market,” *Review, Federal Reserve Bank of St. Louis*, 88(1), January/February, pp. 31–56.

Christensen, R. (1997): *Log-Linear Models and Logistic Regression*. Springer, New York.

Ciochetti, B., Deng, Y., Gao, B. and Yao, R., (2002): “The Termination of Commercial Mortgage Contracts through Prepayment and Default: A Proportional Hazard Approach with Competing Risks”, *Journal of Real Estate Economics*, 30(4), pp 595-633.

Ciochetti, B., Deng, Y., Lee, G., Shilling, J. and Yao, R., (2003): “A Proportional Hazards Model of Commercial Mortgage Default with Originator Bias”, *Journal of Real Estate Finance and Economics*, 27(1), pp 5-23.

Clapp, J.M., Goldberg, G.M., Harding, J.P.,and LaCour-Little,, M., (2001): “Movers and Shuckers:Interdependent Prepayment Decisions”, *Real Estate Economics*, 29(4), pp 411-450.

Clauretie, T.M., (1987), “The Impact of Interstate Foreclosure Cost Differences and the Value of Mortgages on Default Rates”, *Journal of the American Real Estate and Urban Economics Association*, 15(5), pp 152-167.

Clauretie, T.M, Herzog, T., (1990): "The Effect of State Foreclosure Laws on Loan Losses: Evidence from the Mortgage Insurance Industry", *Journal of Risk and Insurance*, 56(3), pp221-233

Cohen, J., Cohen, P., West, S. G. and Aiken, L. S., (1983): *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences*, (2nd ed). Hillside, N.J.: Lawrence Erlbaum

Coles, A. (1992): Causes and Characteristics of Arrears and Possessions, *Housing Finance*, 13(1)., pp 10-12

Collett, D. (1994): *Modelling Survival Data in Medical Research*, Chapman & Hall

Collin-Dufresne, P., Harding, J.P., (1999): A Closed Form Formula for Valuing Mortgages. *Journal of Real Estate Finance and Economics*, 19(2), pp 133-146.

Cooper, A., Meen, G., (2001): *The Relationship Between Mortgage Possessions and the Economic Cycle*, Oxford Economic Forecasting Report to the Association of British Insurers.

Corgel, J.B., Gay, G.D.,(1987): Local Economic Base, Geographic Diversification and Risk Management of Mortgage Portfolios, *AREUEA Journal*, 15(3)., pp256-267

Council of Mortgage Lenders (CML), (2006): *U.K. Mortgage Underwriting*, Report by the CML. Available from http://www.cml.org.uk/cml/media/press/

Council of Mortgage Lenders (CML), (2010): *The outlook for mortgage funding markets in the UK in 2010 – 2015*, Report by the CML. Available from http://www.cml.org.uk/cml/media/press/

Cowan, A., Cowan, C., (2004): Default correlation: An empirical investigation of a subprime lender. *Journal of Banking and Finance*, 28(4), 753–771.

Cox, D. R., (1972): 'Regression models and life-tables', *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2), pp 187-220.

Cox, D. R., Oakes, D., (1984), *Analysis of Survival Data*,Monographs on Statistics and Applied Probability, Chapman and Hall, London.

Cox, J. C., Ingersoll, J. E. Jr., and Ross, S. A., (1979): "Duration and the Measurement of Basis Risk", *Journal of Business,* 52(1), pp 51-61.

Cox, J. C., Ingersoll, J. E. Jr., and Ross, S. A., (1981): "A Re-Examination of the Traditional Hypothesis about the Term Structure of Interest Rates", *Journal of Finance*, 36(4), pp 769-799

Cox, J. C., Ingersoll, J. E. Jr., and Ross, S. A., (1985a): "An Intertemporal General Equilibrium Model of Asset Prices", *Econometrica*, 53(2), pp 363-384

Cox, J. C., Ingersoll, J. E. Jr., and Ross, S. A., (1985b): "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53(2), pp 385-407

Cox, J.C., Ross, S.A., and Rubinstein, M.,(1979): Option Pricing: a Simplified Approach., *Journal of Financial Economics*. 7(2), pp 229-264.

Crank, J. (1984): *Free and Moving Boundary Problems*. Clarendon Press, Oxford.

Crews-Cutts, A., Merrill, W.A.,(2008): "Interventions in Mortgage Default: Policies and Practices to Prevent Home Loss and Lower Costs.", Freddie Mac Working Paper 08-01.

Crews-Cutts A., Van Order, R.,(2005): On the economics of subprime lending. *Journal of Real Estate Finance and Economics*, 30(2), pp 167–196.

Crook, J. N., Edelman, D. B., (Eds), *Credit Scoring and Credit Control*. OUP: Oxford, pp 109-121.

Crook, J. N., Banasik, J., (2004): Does reject inference really improve the performance of application scoring models? *Journal of Banking and Finance,* Vol.28, pp 857–874.

Crook, J. N., Bellotti, T., (2010): Time varying and dynamic models for default risk in consumer loans. *Journal of the Royal Statistical Society series A (Statistics in Society)*, 173(2), pp 279 - 468

Crook, J. N., Bellotti, T., (2012): Asset correlations for credit card defaults, *Applied Financial Economics*, Vol. 22, pp 87-95..

Crook, J. N., Leow, M. and Bellotti, T., (2011): 'Survival Analysis for credit default with unobserved heterogeneity' Credit Research Centre Working Paper.

Crouhy, M., Galai, D., and Mark, R., (2000): "A Comparative Analysis of Current Credit Risk Models", *Journal of Banking and Finance*, 24(1), pp57-117

Cunningham, D., Capone, C.J., (1990): "The Relative Termination Experience of Adjustable to Fixed Rate Mortgages", *Journal of Finance*, 45(5), pp 1687-1703.

Cunningham, D., Hendershott, P. H., (1984): "Pricing FHA Mortgage Default Insurance", *Housing Finance Review*, 3(4), pp 383-392.

Cyert, R.M., Davidson, H.J. and Thomson, G.L. (1962): Estimation of the Allowance for Doubtful Accounts by Markov Chains, *Management Science*, pp287-303

Danis, M.A., Pennington-Cross, A.,(2008): The delinquency of subprime mortgages, *Journal of Economics and Business,* 60(1/2)., pp 67–90

Davidson, A.S., Herskovitz, M.D. and Van Drunen, L.D.,(1988): The Refinancing Threshold Pricing Model: An Economic approach to Valuing MBS, *Journal of Real Estate finance and Economics*, 1(1)., pp 117-130

De Leonardis, D., Rocci, R., (2008): Assessing the Default Risk by means of a discrete-time survival analysis approach, *Applied Stochastic Models in Business and Industry*. 24(2)., pp 291–306

Demyanyk, Y. and Van Hemert, O. Forthcoming. “Understanding the Subprime Mortgage Crisis.” *Review of Financial Studies*. Available at SSRN: http://www.ssrn.com/abstract_1020396

Deng, Y., (1997), “Mortgage Terminations: An Empirical Hazard Model with Stochastic Term Structure”, *Journal of Real Estate Finance and Economics*, 14(3)., pp 309-331.

Deng, Y., Quigley, J.M., and Van Order, R., (1994): Household Income, Equity, and Mortgage Default Risks, University of California-Berkeley, Working Paper.

Deng, Y, Quigley, J.M, and Van Order, R. (1996): Mortgage default and Low Down-payment Loans: The Cost of Public Subsidy, *Regional Science and Urban Economics* 26, pp263-285.

Deng, Y., Quigley, J.M. and Van Order, R., (2000), “ Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options”, *Econometrica*, 68(2)., pp 275-307.

Deng, Y., Quigley, J.M., (2002), Woodhead Behaviour and the Pricing of Residential Mortgages, Working Paper, IBER, University of California, Berkley

Deng, Y., Gabriel, S., (2006): Risk-based Pricing and the Enhancement of Mortgage Credit Availability among Underserved and higher Credit-Risk Populations, *Journal of Money, Credit and Banking*, 38(6)., pp1431-1460

Diaz-Serrano, L.,(2005): Income Volatility and Residential Mortgage Delinquency across the EU, Journal of housing Economics,14(1)., pp153-177

Diggle, P.J., Liang, K.Y., and Zeger, S.L., (1994), *Analysis of Longitudinal Data*, Clarendon Press, Oxford

Dixit, A. K., (1992): "Investment and Hysteresis", *The Journal of Economic Perspectives*, Vol. 6, pp 107-132

Dixit, A. K., Pindyck, R. S.,(1994), Investment under Uncertainty, Princeton University Press, New Jersey.

Douetil, D.J., (1994): "The Inter relationship between the Mortgage and Insurance Industries in the United Kingdom", *Housing Policy Debate*, 5(3), pp 275-306.

Downing, C. (1998): Topics in the Home Mortgage Industry, Dissertation, U. C. Berkeley.

Downing, C., Stanton, R., Wallace, N., (2005): An Empirical Test of a Two-Factor Mortgage Valuation Model: How Much Do House Prices Matter?, *Real Estate Economics*, 33(6), pp 681-710.

Duffie, D., (1996), Dynamic Asset Pricing Theory, 2nd ed., Princeton University Press, Princeton.

Dunn K. B., McConnell, J.J., (1981a), "A Comparison of Alternative Models for Pricing GNMA Mortgage-Backed Securities", *Journal of Finance*, 36(2), pp 471-484.

Dunn K. B., McConnell, J.J., (1981b), "Valuation of GNMA Mortgage-Backed Securities", *Journal of Finance*, 36(3), pp 599-617.

Dunn, K.B., Spatt, C., (2005): The Effect of Refinancing Cost and Market Imperfections on the Optimal Call Strategy and the Pricing of Debt Contracts. *Real Estate Economics*., 33(5), pp595-618.

Efron, B.,(1977): The efficiency of Cox's likelihood function for censored data, *Journal of the American Statistical Association*, 72(5), pp557–565.

Efron, B., (1988): Logistic Regression, Survival Analysis, and the Kaplan Meier Curve, *Journal of the American Statistical Association*, 83(402)., pp414-425

Einstein, A., (1906),"Zur Theorie der Brownschen Bewegung", *Annalen der Physik*, Vol. 19., pp 371-381

Eisenbeis, R. A., (1977): 'Pitfalls in the application of Discriminant Analysis in business, finance and economics', *Journal of Finance*, 32(3), 875-900.

Elbers, C., Ridder, G., (1982): True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model, *Review of Economic Studies*, 49(3)., pp403-409

Elul, R., (2006): Residential Mortgage Default, *Business Review*, Q3, pp21-30

Elul, R., Souleles, N.S., Chomsisengphet, S., Glennon, D and Hunt, R., (2010): What "Triggers" Mortgage Default?, Federal Reserve Bank of Philadelphia WORKING PAPER NO. 10-13

Elmer, P.J., (1997): A Choice-Theoretic Model of Single-Family Mortgage Default, Federal Deposit Insurance Corporation, FDIC Working Paper 97-1

Elmer, P.J. Seelig, S.A.,(1998): The Rising Long Term Trend of Single-Family Mortgage Foreclosure Rates, Federal Deposit Insurance Corporation, FDIC Working Paper 98-2

Elmer, P., Seelig. S.A.,(1999). “Insolvency, Trigger Events, and Consumer Risk Posture in the Theory of Single-Family Mortgage Default,” *Journal of Housing Research*, 10(1), pp. 1-25.

Englund, P., Gordon, T. M., and Quigley, J. M. (1999), “The Valuation of Real Capital: A Random Walk Down Kungsgaten,” *Journal of Housing Economics*, 8(3), pp 205-216.

Epperson, J.F., Kau, J. B., Keenan, D.C., and Muller III, W.J., (1985): “Pricing Default Risk in Mortgages”, *Journal of the American Real Estate and Urban Economics Association*, 13(3), pp 152-167.

Evans, R D., Maris, B A., Weinstein, R L., (1985): Expected Loss and Mortgage Default Risk, *Quarterly Journal of Business & Economics*, 24(1), pp 75-92

Figueira, C., Glen, J. and Nellis, J.,(2005): A Dynamic Analysis of Mortgage Arrears in the UK Housing Market, *Urban Studies*, 42(10)., pp 1755-1769

Findlay, M .C., Capozza, D.R., (1977): “The Variable Rate Mortgage and Risk in the Mortgage Market: An Option Theory Perspective”, *Journal of Money, Credit and Banking*, 9(2), pp 356-364.

Flinn, C.J., Heckman, J.J., (1982): New Methods for Analysing Individual Event Histories, Sociological Methodology, 13(1)., pp 99-140

Follain, J.R., Oldrich, J. and Sinha, G.P.: (1992), Unobserved Transactions Costs and Multifamily Prepayments. University of Wisconsin-Madison, Working Paper.

Foote, C. L., Gerardi, K.S., Goette, L. and Willen P.S., (2008): “Just the facts: An initial analysis of subprime’s role in the housing crisis,” *Journal of Housing Economics*, 17(4), pp. 291–305.

Foote, C. L, Gerardi, K. S., and Willen, P., (2009): "Negative Equity and Foreclosure: Theory and Evidence," *Journal of Urban Economics*, 64(2), pp. 234-245.

Ford, J., Kempson, E., and Wilson, M., (1995): Mortgage Arrears and Possessions: Perspectives from Borrowers, Lenders and the Courts, HMSO.

Ford, J., Burrows, R., Quilgars, D., and Rhodes, D., (2004): Mortgage Insurance: its Role in the Mortgage Safety-Net, University of York, CML Housing Finance, No 62.

Foster, C., Van Order, R., (1984): "An Option Based Model of Mortgage Default", *Housing Finance Review*, 3(4), pp 292-316.

Foster, C., Van Order, R., (1985): "FHA Termination: A Prelude to Rational Option Pricing", *Journal of the American Real Estate and Urban Economics Association*, 13(3), pp 273-291.

Gathergood, J. (2009): 'Income Shocks, Mortgage Repayment Risk and Financial Distress among UK Households', *Working Paper 09/03*, CFCM, School of Economics, University of Nottingham.

Gau, G.W., (1978): A Taxonomic Model for the Risk-Rating of Residential Mortgages, *Journal of Business*, 51(4)., pp687-706

Gerald, C.F., Wheatley, P.O., (5th Ed). 1994, *Applied Numerical Analysis*, Addison-Wesley, Menlo Park, CA.

Gerardi, K. S., Shapiro, A.H. and Willen, P.S., (2008): "Subprime outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures," Federal Reserve Bank of Boston, Working Paper, No. 07-15,

Gerardi, K.S., Shapiro, A.H., and Willen, P.S., (2009): "Decomposing the Foreclosure Crisis: House Price Depreciation versus Bad Underwriting." Federal Reserve Bank of Atlanta, Working Paper 2009-25.

Geske, R., (1979): "The Valuation of Compound Options", *Journal of Financial Economics*, 7(1), pp 63-81.

Gilberto, S.M., Houston, A.L., (1989): "Relocation Opportunities and Mortgage Default", *AREUEA Journal*, 17(1), pp 55-69.

Gilberto, S.M., Ling, D.C., (1992): An Empirical Investigation of the contingent Claims Approach to Pricing Residential Mortgage Debt, *Journal of the American Real Estate and Urban Economics Association*, 20(3), pp 393-426

Goldberg, L., Capone, C.A.,(2002): A Dynamic Double-Trigger Model of Multifamily Mortgage Default, *Real Estate Economics*, 30(1)., pp 85-113

Green, J., Shoven, J.B., (1986): The effects of interest rates on mortgage prepayments, *Journal of Money, Credit, and Banking,* 18(1), pp41-59.

Green, R. K., Shilling, J.D.,(1997): The Impact of Initial Year Discounts on ARM Prepayments, Real Estate Economics, 25(3)., pp 373-385

Green, R. K., LaCour-Little, M.,(1999): Some truths about Ostriches: Who doesn't prepay their mortgages and why they don't, *Journal of Housing Economics,* 8(2), pp 233–48.

Griffith, D.A., Amerhein, C.G., (1997): *Multivariate Statistical Analysis for Geographers*, Prentice Hall, New Jersey.

Guiso, L., Sapienza, P., and Zingales, L., (2010): The determinants of attitudes towards strategic default on mortgages, EUI Working Paper, ECO 2010/31. Available from http://cadmus.eui.eu/bitstream/handle/1814/14296/ECO_2010_31.pdf?sequence=1

Haber, R.J., Schonbucher, P.J., and Wilmott, P.,(1999): Pricing Parisian Options, *Journal of Derivatives.* 6(1) pp 71-79.

Hair, J. F., Anderson, R.L., Anderson, R. E., Tatham, R. L. and Black, W. C., (1998): *Multivariate Data Analysis,* (5th Eds.), Upper Saddle River, N. J.: Prentice-Hall International Inc.

Hall, A.R.,(1985): Valuing The Mortgage Borrowers Prepayment Option, *AREUEA Journal*, 13(3)., pp229-247

Han, A., Hausman, J.A.,(1990): Flexible parametric estimation of duration and competing risk models, *Journal of Applied Econometrics,* 5(1), pp1–28.

Harding, J., Rosenblatt, E., and Yao, V., (2008): "The Contagion Effect of Foreclosed Properties" *Journal of Urban Economics*, Vol. 66, No. 3, pp.164-178

Haughwout, A., Peach, R., and Tracy, J., (2008): "Juvenile Delinquent Mortgages: Bad Credit or Bad Economy?", *Journal of Urban Economics,* 64(2)., pp246–57.

He, H., (1990): "Convergence from Discrete- to Continuous-Time Contingent Claims Prices",*The Review of Financial Studies*, 3(4), pp 523-546.

Heath, D., Jarrow, R.A. and Morton, A., (1991): "Contingent Claim Evaluation with a Random Evolution of Interest Rates," *Review of Futures Markets*, 9(1) pp 54-76.

Heath, D., Jarrow, R.A. and Morton, A., (1992): "Bond pricing and the Term Structure of Interest Rates: A New Methodology For Contingent Claims Evaluation," *Econometrica*, 60(1), pp77-105.

Heckman, J., Singer, B., (1984a): The identifiability of the Proportional Hazard Model, *Review of Economic Studies* 51(2)., pp231–241.

Heckman, J., Singer, B., (1984b): A method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica*, 52(2), pp271–320.

Heckman, J., Singer, B., (1984c): Economic Duration Analysis, *Journal of Econometrics*, 24(84)., pp63-132

Hellebrandt, T, Kawar, S. and Waldron, M., (2009): The Economics and Estimation of Negative Equity, Bank of England Quarterly Bulletin, Q2., pp110-122

Hendershott, P. H., Van Order, R., (1987): "Pricing Mortgages: An Interpretation of the Models and Results", *Journal of Financial Services Research*, 1(1), pp 77-111

Henrici, P.,(1963): *Error Propagation for Difference Methods*, Wiley, New York

Herzog, J. P., Early, J. S., (1970): "Home Mortgage Delinquency and Foreclosure", NBER

Heston, S. L., (1993): "A closed-form solution for options with stochastic volatility with applications to bond and currency options". *The Review of Financial Studies*, 6(2), pp 327-343

Hillegeist, S. A., Keating, E. K., Cram, D.P. and Lundstedt, K.G.,(2004): Assessing Probability of Bankruptcy, Review of Accounting Studies, 9(1)., pp 5-34

Hilliard, J. E., Schwartz, A. L., and Tucker, A. L., (1996), “Bivariate Binomial Options Pricing with Generalised Interest Rate Processes”, *The Journal of Financial Research*, 19(4)., pp 585-602.

Hilliard, J. E., Kau, J. B., and Slawson, V.C., (1998), “Valuing Prepayment and Default in a Fixed Rate Mortgage: A Bivariate Binomial Options Pricing Technique”, *Journal of Real Estate Economics*, 26(3), pp 431-468

Hosmer, D.W., Lemeshow. S., (1989): *Applied Logistic Regression*. Wiley, New York.

Hosmer, D.W., Lemeshow, S., (1999): *Applied Survival Analysis: Regression Modelling of Time to Event Data*, Wiley

Hsu, H.P., (1997): Schaums Outline of Theory and Problems of Probability, Random Variables and Random Processes,McGraw Hill, New York.

Hsu, L., Gorfine, M. and Malone, K.:(2007): On robustness of marginal regression coefficient estimates and hazard functions in multivariate survival analysis of family data when the frailty distribution is mis-specified., *Stat Med*. 26(25): pp4657-78

Hu, J. C., (1988): Derivative Mortgage Securities: An Overview, *Journal of Real Estate Finance and Economics,* 1(1)., pp 95-115

Huang W., Ondrich, J., (2002): Stay, Pay, or Walk Away: A Hazard Rate Analysis of Federal Housing Administration-Insured Multifamily Mortgage Terminations, *Journal of Housing Research*, 13(1), pp 185-117.

Hugonnier, J., (1999): The Feynman-Kac Formula and Pricing Occupation Time Derivatives, *International Journal of Theoretical and Applied Finance,* 2(1), pp153-178.

Hull, J.C., (2006), "Options, Futures and Other Derivatives", 6th Ed, Pearson Prentice Hall, Upper Saddle River, New Jersey.

Hull, J., White, A., (1996), *Hull-White on Derivatives*, Risk Publications, London.

Hurst, E., Stafford, F.,(2004): Home is where the Equity Is: Mortgage Refinancing and Household Consumption, *Journal of Money,Credit and Banking*, 36(6).,
pp 985-1014

Ito, K. (1951): On Stochastic Differential Equations. Memoirs, *American Mathematical Society*, 4(1), pp 1-51.

Jackson, J. J., Kaserman, D. L., (1980): "Default Risk on Home Mortgage Loans: A test of Competing Hypotheses", *Journal of Risk and Insurance*, 47(4)., pp 678-690.

Jarrow, R.A. (1988), "Finance Theory", Prentice Hall International, Englewood Cliffs, New Jersey.

Jondeau, E., Poon, S., Rockinger, M., (2006), *Financial Modelling Under Non-Gaussian Distributions*, Springer Verlag, New York.

Jones, M.P. (1996). Indicator and stratification methods for missing explanatory variables in multiple linear regressions, *Journal of the American Statistical Association*, Vol. 91, pp222-230

Joseph, M.(2007): "Attaining sharper Pricing to risk in the UK Subprime Mortgage Market" from the Conference proceedings of the Credit Scoring and Credit Control conference. University of Edinburgh.

Jung, A., (1962): Terms on Conventional Mortgage Loans on Existing Homes, *Journal of Finance*, pp 432-443.

Kalbfleisch, J.D. Prentice, R.L., (1980) *The Statistical Analysis of Failure Time Data*, Wiley, New York

Kalotay, A.J., Yang, D. and Fabozzi, F. J., (2004): "An Option-theoretic Prepayment Model for Mortgages and Mortgage-backed Securities", *International Journal of Theoretical and Applied Finance*, 7(8)., pp. 949-978

Kariya, T., Pliska, S. R., and Ushiyama, F., (2002): A 3-factor valuation model for mortgage backed securities (MBS), Working Paper, Kyoto Institute of Economic Research.

Kao, E. P.C., (2006), An Introduction to Stochastic Processes, Thomson, New York.

Kaplan, E. L., Meier, P., (1958): Nonparametric estimation from incomplete observations, *Journal of American Statistical Association*, 53, pp 457–481.

Karlin, S., Taylor, H.M., (1975), A First Course in Stochastic Processes, 2^{nd} Ed, Academic Press, London.

Kau, J. B., Keenan, D.C., (1980), "The Theory of Interest Rates and Housing", *Journal of Financial and Quantitiative Analysis*, 15(4), pp 833-847.

Kau, J. B., Keenan, D.C., (1981), "On The Theory of Interest Rates, Consumer Durables and the Demand for Housing", *Journal of Urban Economics*, 10(3), pp 320-331.

Kau, J. B., Keenan, D.C., (1983), "Inflation, Taxes, and Housing in a Growing Economy", *Journal of Public Economics*, 21(1), pp 93-104.

Kau, J. B., Keenan, D.C., (1995): "An Overview of Option Theoretic Pricing of Mortgages", *Journal of Housing Research*, 6(2), pp 217-244.

Kau, J. B., Keenan, D.C., (1999), "Patterns of Rational Default", *Regional Science and Urban Economics*, 29(7), pp 765-785.

Kau, J. B., Keenan, D.C., (2000): Catastrophic Default and Credit Risk for Lending Institutions, *Journal of Financial Services Research*,15(2), pp 87-102.

Kau, J. B., Keenan, D.C., and Kim, T., (1993): Transaction Costs, Suboptimal Termination and Default Probabilities, *Journal of the American Real Estate and Urban Economics Association*, 21(3), pp 247-263.

Kau, J. B., Keenan, D.C., and Kim, T., (1994): "Default Probabilities for Mortgages", *Journal of Urban Economics*, 35(3), pp 278-296.

Kau, J. B., Kim, T., (1993): The Timing of Prepayment: A Theoretical Analysis, *Journal of Real Estate Finance and Economics*, 7(2)., pp 221-228

Kau, J. B., Kim, T., (1994): "Waiting to Default: The Value of Delay", *Journal of the American Real Estate and Urban Economics Association*, 22(3), pp 539-551.

Kau, J. B., Keenan, D.C., Muller III, W.J., and Epperson, J.F., (1985): "Rational Pricing of Adjustable Rate Mortgages", *Journal of the American Real Estate and Urban Economics Association*, 13(2), pp 117-128.

Kau, J. B., Keenan, D.C., Muller III, W.J., and Epperson, J.F., (1987): "The Valuation and Securitisation of Commercial and Multifamily Mortgages" , *Journal of Banking and Finance*, 11(5), pp 525-546.

Kau, J. B., Keenan, D.C., Muller III, W.J., and Epperson, J.F., (1990): "The Valuation and Analysis of Adjustable Rate Mortgages", *Management Science*, 36(12), pp 1417-1431.

Kau, J. B., Keenan, D.C., Muller III, W.J., and Epperson, J.F., (1992): "A Generalized Valuation Model for Fixed Rate Residential Mortgages ", *Journal of Money, Credit and Banking*, 24(3), pp 279-299.

Kau, J. B., Keenan, D.C., Muller III, W.J., and Epperson, J.F., (1993): "Option Theory and Floating Rate Securities with a Comparison of Adjustable and Fixed Rate Securities", *Journal of Business*, 66(4), pp 595-618.

Kau, J. B., Keenan, D.C., Muller III, W.J., and Epperson, J.F., (1995): "The Valuation at Origination of Fixed Rate Motgages with Default and Prepayment", *Journal Real Estate Finance and Economics*, 11(1) pp 5-36.

Kau, J. B., Keenan, D.C., and Smurov, A.A., (2006): Reduced Form Mortgage Pricing as an Alternative to Option Pricing Models, *Journal of Real Estate Finance and Economics*, 33(1)., pp 183-196

Kau, J.B,, Keenan, D.C,, and Li, X., (2011): An Analysis of Mortgage Termination Risks: A Shared Frailty Approach with MSA-Level Random Effects, *The Journal of Real Estate Finance and Economics*, 42(1), pp 51-67.

Kau, J. B., Keenan, D. C., Lyubimov, C., and Slawson, V. C., (2011): Subprime Mortgage Default, *Journal of Urban Economics*, 70(2), pp 75-87

Kau, J. B, Keenan, D. C., Lyubimov, C., and Slawson, V. C., (2012): "Asymmetric Information in the Subprime Mortgage Market," *The Journal of Real Estate Finance and Economics*, 44(1), pp 67-89

Kau, J. B,, Keenan, D.C., and Smurov, A.A., (2011): Leverage and Mortgage Foreclosures, *The Journal of Real Estate Finance and Economics*, 42(4), pp 393-415.

Kau, J. B., Springer, T.M., (1992a): The Prepayment of fixed-Rate Mortgages Underlying Mortgage-Backed Securities: A Switching Regimes Analysis, *Journal of Housing Research*, 2(2)., pp 257-275

Kau, J. B., Springer, T.M., (1992b): An Analysis of Financial and Nonfinancial Prepayment of GNMA Securities with a Varying Coefficient Model, *Journal of Real Estate Research,* 8(1)., pp69-86

Kau, J. B., Slawson, V.C., (2002): "Frictions, heterogeneity and optimality in mortgage modelling". *Journal of real estate finance and economics,* 24(3), pp239-260

Kau, J. B, Keenan, D. C., and Yildirim, Y., (2009): Estimating Default Probabilities Implicit in Commercial Mortgage Backed Securities (CMBS), *The Journal of Real Estate Finance and Economics,* 39(2), pp 107-117

Kelly, A., Slawson, V.C., (2001): "Time-varying mortgage prepayment penalties". *Journal of real estate finance and economics* 23(2) : pp235-254

Kim, T., (1987), " A Contingent Claims Analysis of Price Level-Adjusted Mortgages", *Journal of the American Real Estate and Urban Economics Association*, 15(3)., pp 117-131.

Klein, J.P., Moeschberger, M.L., (1997) *Survival Analysis: Techniques for Censored and Truncated Data*, Springer Verlag, Inc., New York

Kuo, C-L., (1996),"Housing Price Dynamics and the Valuation of Mortgage Default Options", *Journal of Housing Economics*, 5(1)., pp 18-40.

Kwok, Y., Lau. K., (2001): Pricing Algorithms for Options with Exotic Path-Dependence. *Journal of Derivatives*, 9(1), pp 28-38.

LaCour-Little, M.,(1999): Another Look at the Role of Borrower Characteristics in Predicting Mortgage Prepayment, *Journal of Housing Research*, 10(1)., pp 45-60

LaCour-Little, M., Chun, G.H.,(1999): Third Party Originators and mortgage Prepayment Risk: An Agency Problem?, *Journal of Real Estate Research*, 17(1/2)., pp55-70

LaCour-Little, M.,(2004): Equity Dilution: An Alternative Perspective on Mortgage Default, *Real Estate Economics*, 32(3)., pp 359-384

Lambrecht, B, Perraudin, W, and Satchell, S., (1997): "Time to Default in the UK Mortgage Market", *Economic Modelling*, 14(4)., pp 485-499

Lambrecht, B., Perraudin, W., and Satchell, S., (2003): 'Mortgage Default and Possession Under Recourse: A Competing Hazards Approach', *Journal of Money, Credit, and Banking,* 35(3), pp. 425-442.

Lancaster, T.: (1990), *The Econometric Analysis of Transition Data*. pp 154-155, New York, Cambridge University Press.

Land, K.C., Nagin, D.S. and McCall, P.L. (2001), Discrete Time Hazard Regression Models with Hidden Heterogeneity: The Semiparametric Mixed Poisson Regression Approach, *Sociological Methods and Research*, 29(3)., pp 342-373

Lawrence, E. C., Smith, L.D. and Rhoades, M.,(1992): An analysis of default risk in mobile home credit, *Journal of Banking and Finance*, 16(2), 299-312.

LeClere, M. J. (2000), 'The occurrence and timing of events: Survival analysis applied to the study of financial distress', *Journal of Accounting Literature*, 19(1), pp 158-189.

Leece, D., (2004): *Economics of the Mortgage Market: Perspectives on Household Decision Making.* Blackwell Publishing Ltd.

Lekkas, V., Quigley, J and Van Order, R., (1993), "Loan Loss Severity and Optimal Mortgage Default", *Journal of the American Real Estate and Urban Economics Association*, 21(4)., pp 353-372.

Leow, M., (2011): 'Competing Risks Survival Model for Mortgage Loans with Simulated Loss Distributions. Credit Research Centre, Working Paper. University of Edinburgh.

Leow, M., Mues, C., (2012): Predicting loss given default (LGD) for residential mortgage loans: A two-stage model and empirical evidence for UK bank data. *International Journal of Forecasting*., 28(1), pp 183–195

Leung, W.K., Sirmans, C.F., (1990): "A Lattice Approach to Pricing Fixed Rate Mortgages with Default and Prepayment Options", *Journal of the American Real Estate and Urban Economics Association*, 18(1), pp 91-104.

Liao, S., Tsai, M. and Chiang, S.,(2008): Closed Form Mortgage Valuation Using a Reduced Form Model, *Real Estate Economics*, 36(2)., pp 313-347

Longstaff, F. A., (2004): Optimal recursive refinancing and the valuation of mortgage-backed securities, NBER Working Papers: 10422.

Longstaff F.A., (2005): Borrower Credit and the Valuation of Mortgage-Backed Securities. *Real Estate Economics*, 33(6), pp 619-661.

Longstaff, F. A., Schwartz, E.S.,(1995): A simple approach to valuing risky fixed and floating rate debt, *Journal of Finance* 50(6), pp789–819.

Marschak, J., (1960); Binary choice constraints on random utility indications, in K. Arrow, ed., Stanford Symposium on Mathematical Methods in the Social Sciences, 312–329 (Stanford University Press, Stanford, CA).

Mattey, J., Wallace, N.,(2001):Housing-Price Cycles and Prepayment Rates of US Mortgage Pools, *Journal of Real Estate Finance and Economics*, 23(2)., pp161-184

Maxam, C.L., LaCour-Little, M., (2001): Applied Nonparametric Regression Techniques: Estimating Prepayments on Fixed Rate Mortgage-Backed Securities, *Journal of Real Estate Finance and Economics*, 23(2)., pp139-161

May, O., Tudela, M., (2005): When is mortgage Indebtedness a Financial Burden to British Households? A Dynamic Probit Approach, Bank of England Working Paper 277

Mayer, C.J., Pence, K., (2008): “Subprime Mortgages: What, Where, and to Whom?” Federal Reserve Board of Governors Finance and Economic Discussion Series Paper 2008-29.

Mayer, C.J., Pence, K. and Sherlund, S., (2009): The Rise in Mortgage Defaults, *Journal of Economic Perspectives*, 23(1), pp 27-50.

McConnell, J.J., Singh, M., (1994): “Rational prepayments and the valuation of collateralized mortgage obligations”. *Journal of finance* 49(3): pp 891-92

McFadden, D., (1973): "Conditional Logit Analysis of Qualitative Choice Behaviour," in P. Zarembka (Ed.), *Frontiers in Econometrics*, Academic Press, New York

McFadden, D., (1984): Econometric Analysis of Qualitative Response Models, in Z. Griliches, and M. D. Intriligator, eds., *Handbook of Econometrics, volume II*, pp 1396–1457 (Elsevier Science, New York, NY).

McFadden, D., and Train, K., (2000): Mixed MNL Models for discrete response, *Journal of Applied Econometrics* 15(4), pp 447–470.

Meese, R.A., Wallace, N.E., (1994): Testing the Present Value Relation for House Prices: Should I Leave my House in San Francisco? *Journal of Urban Economics*, 35(3), pp 245-66.

Meese, R.A., Wallace, N.E. (1997): The Construction of Residential Housing Price Indices: A comparison of Repeat-Sales, Hedonic Regression and Hybrid Approaches, *Journal of Real Estate Finance and Economics* 14(1): pp51-73

Meese R.A., Wallace, N. E., (1999): Dwelling Price Dynamics in Paris, France Working Paper No. W98-003, Institute of Business and Economic Research - Available at http://urbanpolicy.berkeley.edu/pdf/meese_wallace.pdf

Merton. R. C., (1971): "Optimum consumption and portfolio rules in a continuous time model". *Journal of Economic Theory*, 3(3). pp 373-413

Merton, R, C., (1973): "Theory of Rational Option Pricing", *The Bell Journal of Economics and Management*, 4(1), pp 141-183.

Merton, R., (1974): "On the pricing of corporate debt: the risk structure of interest rates", *Journal of Finance*, 29(2), pp 449-470

Merton, R., (1977): "An analytic derivation of the cost of deposit insurance and loan guarantees", *Journal of Business and Banking*, 1(1), pp 3-11

Mian, A., Sufi, A., (2009): The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis, *Quarterly Journal of Economics*, 124(4).

Mian, A., Sufi, A., (2010): "House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis" Chicago Booth Research Paper No. 09-20.

Mickey, J., Greenland. S., (1989): A study of the impact of confounder selection criteria on effect estimation. *American Journal of Epidemiology*, Vol. 129, pp 125-137.

Miles, D., (2003): *The U.K. Mortgage Market: Taking a Longer-Term View, Interim Report: Information, Incentives, and Pricing*. HM Treasury.

Mitropoulos, A., Zaida, R. (2009): Relative Indicators of Mortgage Default risk among UK residential Mortgages, MPRA Paper 19619 Available Online at http://mpra.ub.uni-muechen.de/19619/

Moger, T.A., Aalen, O. O.,(2005): A distribution for multivariate frailty based on the compound Poisson distribution with random scale., *Lifetime Data Anal*., 11(1), pp 41-59.

Mood, A.M., Graybill, F.A. and Boes, D.C.: (1974), Introduction to the theory of Statistics, International Ed., McGraw Hill, Singapore

Morrison, D.F., *Multivariate Statistical Methods* (2nd Ed). *1976*. McGraw-Hill, New York.

Morton, T.G., (1975): "A Discriminant Function Analysis of Residential Mortgage Delinquency and Foreclosure", *AREUEA Journal*, 3(1), pp 73-90.

Morton, K. W., and Myers, D.F, (1994): "Numerical Solutions of Partial Differential Equations", Cambridge University Press, Cambridge.

Muellbauer, J., Cameron, G., (1997): 'A Regional Analysis of Mortgage Possessions: Causes, Trends and Future Prospects', *Housing Finance*, 34(1)**,** pp. 25–34.

Muellbauer, J., Murphy, A., (1997): 'Booms and Busts in the UK Housing Market', *Economic Journal*, 107(445), pp1701-1727

Munro, M., (2000): Labour-market insecurity and risk in the owner-occupation housing market, *Environment & Planning* A, Vol. 32, pp1375-1389.

Narain, B., (1992): Survival Analysis and the credit granting decision. In: Thomas L. C, Crook, J. N., Edelman, D. B. (Eds). *Credit Scoring and Credit Control*, OUP: Oxford, pp 109-121.

Navratil, F J.,(1985), "The Estimation of Mortgage Prepayment Rates," *Journal of Financial Research*, 8(2), pp107-117.

Neftci, S.N., (2000), *An Introduction to the Mathematics of Financial Derivatives*, 2nd Ed, Elsevier, Academic Press, San Diego, California.

Nelson, D.B., Ramaswamy, K., (1990), "Simple Binomial Processes as Diffusion Approximations in Financial Models", *The Review of Financial Studies*, 3(3), pp 393-430.

Nur, U., Shack, L.G., Rachet, B., Carpenter, J.R. and Coleman, M.P., (2010): Modelling relative survival in the presence of incomplete data: a tutorial, *International Journal of Epidemiology*, Vol.39, pp118–128

Oksendal, B., (1995), *Stochastic Differential Equations: An Introduction with Applications*, 4th Ed, Springer Verlag, New York.

Page, A., (1964): The Variation of Mortgage Interest Rates, *Journal of Business*, pp 280-294

Pavlov, A. D., (2001), "Competing Risks of Mortgage Prepayments: Who Refinances, Who Moves and Who Defaults?", *Journal of Real Estate Finance and Economics*, 23(2)., pp 185-211

Pennington-Cross, A., (2002): Subprime lending in the primary and secondary markets, *Journal of Housing Research*, *13*(1)., pp31–50.

Pennington-Cross, A., (2003): Credit history and the Performance of Prime and nonprime Mortgages, *Journal of Real Estate Finance and Economics*, 27(3)., pp 279-301

Pennington-Cross, A., (2006): The duration of foreclosures in the subprime mortgage market: A competing risks model with mixing, Federal Reserve Bank of St. Louis, Working Papers: 2006-027.

Pennington-Cross, A., (2010): The Duration of Foreclosures in the Subprime Mortgage Market: A Competing Risks Model with Mixing, *Journal of Real Estate Finance and Economics*, 40(1)., pp109-129

Pennington-Cross, A., Ho, G., (2006): Termination of Subprime Hybrid and Fixed Rate Mortgages, Federal Reserve Bank of St .Louis, Working Paper July 2006

Pennington-Cross, A., Ho, G., (2010): The Termination of Subprime Hybrid and Fixed Rate Mortgages. *Real Estate Economics*, 38(3)., pp399-426.

Peristiani, S., Bennett P, Monsen G, Peach, R. and Raiff, J., (1997): Effects of household creditworthiness on mortgage refinancing, Federal Reserve Bank of New York Economic Policy Review, 83-103.

Phillips, R. A., Rosenblatt, E. and Vanderhoff, J.H.,(1996), The Probability of Fixed and Adjustable- Rate Mortgage Termination, *Journal of Real Estate Finance and Economics*, 13(1), pp 95-104.

Pierce, D.A., Stewart, W.H., and Kopecky, K.J., (1979), Distribution-free Regression Analysis of Grouped Survival Data, *Biometrics*, 35, pp785-793.

Pindyck, R., (1991): "Irreversibility, Uncertainty and Investment", Journal of Economic Literature, Vol. 29, pp 1110-1148.

Posey, L. L., Yavas, A., (2001), "Adjustable and Fixed Rate Mortgages as a Screening Mechanism for Default Risk", *Journal of Urban Economics*, 49(1)., pp 54-79.

Prentice, R. L., Williams, B. J., and Peterson, A.V. (1981): On the Regression Analysis of Multivariate Failure Time Data, *Biometrika*, 68(3), pp373-379

Prentice, R.L., Gloeckler, L.A.,(1978), Regression Analysis of Grouped Survival Data with Applications to Breast Cancer, *Biometrics*, 34(1), pp57-67.

Press, W., Teukolsky, S.,Vetterling, W., and Flamery, B., (1992). *Numerical Recipes in C, The Art of Scientific Computing* (2nd Ed.). Cambridge: Cambridge University Press.

Qi, M. Yang, X., (2009): Loss Given Default of High Loan to Value Residential Mortgages, *Journal of Banking and Finance*, Vol. 33, pp788-799.

Quercia, R.G., Stegman, M.A.,(1992): Residential Mortgage default : A review of the literature, *Journal of Housing Research*, 3(2), pp 341-379

Quercia, R. G., Stegman, M.A. and Davis, W.R., (2005): The Impact of Predatory Loan Terms on Subprime Foreclosures: The Special Case of Prepayment Penalties and Balloon Payments, Working Paper January 2005, Centre for Community Capitalism, University of North Carolina.

Quigley, J.M., (1987): Interest Rate Variations, Mortgage Prepayments and Household Mobility, *The Review of Economics and Statistics*, pp636-643

Quigley, J. M., Van Order, R., (1990): Efficiency in the Mortgage Market: The Borrower's Perspective, *AREUEA Journal*, 18(3)., pp237-253

Quigley, J. M., Van Order, R., (1991): Defaults on mortgage obligations and capital requirements for US savings institutions A policy perspective, *Journal of Public Economics*, 44(2)., pp 353-369

Quigley, J. M., Van Order, R., (1995): "Explicit Tests of Contingent Claims Models of Mortgage Default", *Journal of Real Estate Finance and Economics*, 11(1), pp 99-117.

Quigley, J.M., Van Order, R. and Deng, Y., (1994),:" The Competing Risks for Mortgage Termination by Default and Prepayment: A Minimum Distance Estimator", Working Paper, NBER, Cambridge, MA

Rebonato, R., (1998): Interest Rate Option Models, 2nd Ed, John Wiley & Sons, Chichester.

Rendlemann, R., Bartter, B., (1979): Two-State Option Pricing. *Journal of Finance*, 34, pp 1093-1110.

Rice, J.A., (1995), Mathematical Statistics and Data Analysis, 2nd Ed, Duxbury Press, California

Rose, M. J.,(2008): Predatory Lending Practices and Subprime Foreclosures: Distinguishing Impacts by Loan Category, *Journal of Economics and Business,* 60(1/2), pp 13-32.

Sa-Aadu, J., (1988): "Legal Restrictions, Credit Allocation and Default Risk under Fixed and Adjustable Rate Mortgages", *Housing Finance Review*, 7(2), pp 225-247.

Sa-Aadu, J., Megbolugbe, I., (1995): Heterogeneous Borrowers, Mortgage Selection and Mortgage Pricing, *Journal of Housing Research*, 5(2), pp 333-348.

Sa-Aadu, J., Sirmans, C. F., (1995): Differentiated Contracts, Heterogeneous Borrowers, and the Mortgage Choice Decision, *Journal of Money, Credit and Banking*, 27(2), pp 498-510.

Samuelson, P.: (1973), Mathematics of speculative price, *SIAM Review*, 15(1), pp1-42

Sandor, R.L., Sosin, H.B., (1975): "The Determinants of Mortgage Risk Premiums: A Case study of the Portfolio of a Savings and Loans Association", *Journal of Business* 48(1), pp27-38

Saunders, M., Lewis, P., and Thornhill, A., 2nd Ed (2000): *Research Methods for Business Students*, Pearson, Essex.

Schurz, H.,(2001): Brief Introduction to Numerical Analysis of Stochastic Differential Equations without tears in Kannan, D and Lakshmikantham, V, 2001: *Handbook of Stochastic Analysis and Applications*, pp 239-361, Dekker, Basel.

Schwartz, E. S., Torous, W. N., (1989), "Prepayment and the Valuation of Mortgage Backed Securities", *The Journal of Finance,* 44(2)., pp 375-392

Schwartz, E. S., Torous, W. N., (1991): "Caps on Adjustable Rate Mortgages: Valuation Insurance and Hedging", In Financial Markets and Financial Crises, Ed. R. G. Hubbard. University of Chicago Press, Chicago.

Schwartz, E.S., Torous, W.N., (1992): Prepayment, Default and the Valuation of Mortgage Pass-through Securities. *Journal of Business*, 65(2)., pp 221-239.

Schwartz, E.S., Torous, W. N.,(1993): Mortgage Prepayment and Default Decisions: A Poisson Regression Approach. *Journal of the American Real Estate and Urban Economics Association*, 21(3), pp 431-449.

Schwartz, E. S., Torous W, N., (2003): Mortgage Prepayment and Default Decisions: A Poisson Regression Approach, *Real Estate Economics,* 21(4), pp 431-449.

Seuyoshi, G., (1992): Semi parametric proportional hazard estimation of competing risks model with time-varying covariates, *The Journal of Econometrics* 51(1), pp 25-58.

Sharp, N.J., Newton, D.P., and Duck, P.W., (2008): "An Improved Fixed-Rate Mortgage Valuation Methodology with Interacting Prepayment and Default Options", *Journal of Real Estate Finance and Economics*, 36(3). pp. 307-342.

Sherris, M., (1994): A One-factor Interest Rate Model and the Valuation of loans with Prepayment Provisions, Transactions of the Society of actuaries, 16(2), pp 251-301

Shilling, J.D.,(1995): Rates of Return on Mortgage-Backed Securities and Option Theoretic Models of Mortgage Pricing, Journal of Housing Research, 6(2).,
pp 265-284

Shumway, T., (2001): 'Forecasting bankruptcy more accurately: A simple hazard model', *Journal of Business*, 74(1), pp 101-124.

Singer, J. D., and Willett, J. B.,(1993): It's About Time: Using discrete-time Survival Analysis to Study Duration and the Timing of Events, *Journal of Educational Statistics,* 18(2)., pp 155-195

Singer, J. D., and Willett, J. B.,(2003): *Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence*, Oxford University Press, New York.

Skinner, F. (1999): Modelling Redemption Fees and Incentives on UK Home Mortgages and Modelling Variable and Fixed Rate Lending. Research Paper 18, *Office of Fair Trading*.

Smith C.W. Jr., (1976): "Option Pricing: A Review", *Journal of Financial Economics*, 3(1), pp 3-51.

Smith, G.D. (1978). Numerical *Solution of Partial Differential Equations: Finite Difference Methods*. Clarendon Press, Oxford.

Smith, D., Lawrence, E., (1995): "Forecasting Losses on a Liquidating Long-term Loan Portfolio," *Journal of Banking and Finance*, Vol.19, pp 959-985

Smith, L, Sanchez, S., and Lawrence, E., (1996): "Comprehensive Model for Managing Credit Risk on Home Mortgage Portfolios," *Decision Sciences*, 27(2)
pp. 291-317

Springer, T., Waller, N., (1993a): "Termination of Distressed Residential Mortgages: an Empirical Analysis," *Journal of Real Estate Finance and Economics*, 7(1) pp 43-53

Springer, T., Waller, N., (1993b): "Lender Forbearance: Evidence From Mortgage Delinquency Patterns," *Journal Of The American Real Estate And Urban Economics Association, 21*(1), pp.27-46

Stanton, R., (1995), "Rational Prepayment and the Valuation of Mortgage-Backed Securities", *Review of Financial Studies*, Vol. 8., No. 3., pp 677-708

Stanton, R., Wallace, N., (1995):.ARM Wrestling: Valuing Adjustable Rate Mortgages Indexed to the Eleventh District Cost of Funds. *Real Estate Economics*, 23(2), pp311-345.

Stanton, R., Wallace, N., (1998): "Mortgage Choice: What is the Point?," *Real Estate Economics*, 26(2), pp173-205

Stanton, R., Wallace, N.,(1999): Anatomy of an ARM: The Interest-Rate Risk of Adjustable-Rate Mortgages, *Journal of Real Estate Finance and Economics*., 19(1), pp 49-67

Stepanova, M., Thomas, L.C., (2001): PHAB scores: Proportional hazards analysis behavioural scores. *J. Operational Research Society*, Vol. 52, pp 1007-16.

Stepanova, M., Thomas, L.C., (2002): Survival analysis methods for personal loan data. *Operations Research*, Vol. 50, pp 277-289

Stroebelz, J., (2011): The impact of asymmetric information about collateral values in mortgage lending, Stanford University. Job Market Paper, December 2011. Available from: www.stanford.edu/~stroebel/J_Stroebel_JMP.pdf.

Thomas, L.C., Edelman, D.B., and Crook, J.N., (2002): *Credit Scoring and its Applications*, SIAM, Philadelphia.USA.

Thomas, L. C., Banasik, J., and Crook, J. N., (1999): Not if but when loans default. *J. Opl Res Soc*, 50, pp 1185-1190.

Thurstone, L., (1927): A law of Comparative Judgement, *Psychological Review* Vol 34, pp 273–286.

Tingay, J., Wilkinson, G.: (2002), “The Use of Affordability Data: Does it Add Real Value?”, *Credit Risk International,* May-June.

Titman, S., Torous, W. N., (1989): “Valuing Commercial Mortgages: An Empirical Investigation of the Contingent Claims Approach to Pricing Risky Debt”, *The Journal of Finance*, 4(2), pp 345-373

Turnbull, S.M.,(2005): Unresolved Issues in Modelling Credit-Risky Assets, *The Journal of Fixed Income*, 1, pp 68-87

Vandell, K.D., (1978): “Default risk under alternative mortgage instruments”, *Journal of Finance*, 33(12), pp 1279-1296

Vandell, K. D., (1993): “Handing Over the Keys: A Perspective on Mortgage Default Research, *Journal of the American Real Estate and Urban Economics Association*, Vol. 21, No. 3, pp 211-246

Vandell, K.D. (1995): How Ruthless is Mortgage Default? : A Review and Synthesis of the Evidence. *Journal of Housing Research*, 6(2), pp 245-264.

Vandell, K.D., Thibodeau, T., (1985): “Estimation of Mortgage Defaults using Disaggregate Loan History Data*”, Journal of the American Real Estate and Urban Economics Association*, Vol 13, No.3, pp 292-316.

Vanderhoff, J.,(1996):Adjustable and Fixed Rate Mortgage Termination, Option Values and Local Market Conditions: An Empirical Analysis, *Real Estate Economics*, 24(3)., pp 379-406

Van Order, R.,(2008): Modelling and Evaluating the Credit Risk of Mortgage Loans: A Primer, *Journal of Credit Risk Model Validation*, 2(2)., pp60-82

Vaupel, J.W., Manton, K.G. and Stallard, E., (1979), The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality, *Demography*, 16(3)., pp439-454

Vetzal, K.R., P.A. Forsyth. (1999): Discrete Parisian and Delayed Barrier Options: A General Numerical Approach. *Advances in Futures and Options Research*, 10(1), pp 1-15.

Von Furstenburg, G. M., (1969), "Default Risk on FHA-Insured Home Mortgages as a Function of the Terms of Financing: A Quantitative Analysis, *Journal of Finance*, 24(3), pp 459-477.

Von Furstenburg, G. M., (1970a)," The Investment Quality of Home Mortgages*",* *Journal of Risk and Insurance*, 37(3)., pp 437-445.

Von Furstenburg, G. M., (1970b), "Risk Structures and the Distribution of Benefits within the FHA Home Mortgage Insurance Program", *Journal of Money, Credit and Banking*, 2(3), pp 303-322

Von Furstenburg, G. M., Green, R.J., (1974a): Estimation of Delinquency Risk for Home Mortgage Portfolios, Journal of the American Real Estate and Urban Economics Association, 1, pp 5-19

Von Furstenburg, G. M., Green, R.J., (1974b): Home Mortgage Delinquencies: A Cohort Analysis, *The Journal of Finance,* pp 1545-1548

Wallace, N., (2005): Innovations in Mortgage Modelling, *Real Estate Economics*, 33(4)., pp 587-593

Waller, N.G., (1988): "Residential Mortgage Default: A Clarifying Analysis", *Housing Finance Review*, Vol. 7, pp 321-333.

Weagley, R.O., (1988): Consumer Default of Delinquent Adjustable Rate Mortgage Loans, The *Journal of Consumer Affairs*, 22(1)., pp38-54

White, A. M., (2004): Risk-Based Mortgage Pricing: Present and Future Research, *Housing Policy Debate*, 15(3)., pp 503-531

Whitehead, C. M. E., (1994): The UK Housing Market- Introduction and Update, *Housing Policy Debate*, 5(3)., pp 231-240.

Whitley, J., Windram, R. and Cox, P.,(2004): An Empirical Model of Household Arrears, Bank of England Working Paper 214

Wienke, A., Ripatti, S., Palmgren, J., and Yashin, A.,(2010): A Bivariate survival model with compound Poisson frailty. *Stat Med*., 29(2) pp 275-83.

Willett, J. B., Singer, J. D., (1993): Investigating Onset, Cessation, Relapse and Recovery: Why you should and How you can, use Discrete-Time Survival Analysis, *Journal of Consulting and Clinical Psychology*, 61(6).,952-965

Willett, J. B., Singer, J. D., (1995): It's déjà vu all over again: Using Multiple Spell Discrete-Time Survival Analysis, *Journal of Educational and Behavioural Statistics*, 20(1)., pp 41-67

Williams, A.O., Beranek, W. and Kenkel, J. (1974): Default Risk in Urban Mortgages: A Pittsburgh Prototype Analysis, *AREUEA Journal*, pp 101-112

Wilmott, P., (2001a): *Paul Wilmott Introduces Quantitative Finance*, John Wiley & Sons, Chichester.

Wilmott, P., (2001b): *Paul Wilmott on Quantitative Finance*, John Wiley & Sons, Chichester.

Wilmott, P., (2001c): *Derivatives: The Theory and Practice of Financial Engineering*, John Wiley & Sons, Chichester.

Wilmott, P., Howison, S., and Dewynne, J., (2001d): *The Mathematics of Financial Derivatives: A Student Introduction*, Cambridge University Press, New York.

Yang, T., Buist, H. and Megbolugbe, I., (1999): An Analysis of the Ex Ante Probabilities of Mortgage Prepayment and Default, *Real Estate Economics* 26(4), pp 651–676.

Zipkin, P., (1993), Mortgages and Markov Chains: A simplified valuation model. *Management Science*, 39, pp 683-691.

Zorn, P. M., Lea, M. J., (1989) Mortgage borrower prepayment behaviour: A microeconomic analysis with Canadian adjustable-rate mortgage data, *American Real Estate and Urban Economics Association Journal*, 17(1), pp118-136.

The Homepages of the following Institutions can be accessed at

Academetrics: http://www.acadametrics.co.uk/HousePrices.php

Conference Board: http://www.conference-board.org/data/bciarchive.cfm?cid=2

Council of Mortgage Lenders: http://www.cml.org.uk

Office of National Statistics: http://www.ons.gov.uk/ons/index.html

Appendix A: Profitability of Mortgage Loans

Amid all the changes and controversies in the financial sector, lenders continue to originate loans because the interest received as payment by the borrower over the lifetime of the loan is profitable.

One measure of this profitability is obtained from Return on Capital (ROC).

In a general sense, ROC can be defined as the ratio of Net Income to Capital.

As applied in our instance to a mortgage, the capital is the amount of the loan and the net income is the interest paid.

$$ROC = \frac{\sum PV(Interest\ Payments)}{Capital} = \frac{\sum PV(Payments) - Capital}{Capital}$$

Here PV represents the Present Value, which adjusts for the time value of money.

Since the term of the loan is long (25 years) the cash flows that are generated by the mortgage interest payments are spread out over different economic cycles. It is therefore necessary to take the present value of these cash flows in order to adjust for inflation as the value changes over time. Let us assume the Inflation rate in the calculations that follow is taken to be around 3% per annum on the average as inflation fluctuates. By treating a fixed rate mortgage as an Annuity, the Present Value (PV) may be calculated as

$$PV(Annuity) = m\left[\frac{1-(1+r)^{-n}}{r}\right] \quad \ldots\ \ldots\ \ldots \text{ Equation (A.1)}$$

where, m represents the monthly payment, r represents the inflation rate (expressed monthly) and n represents the number of months in which a payment was received.

Loan Payback Situations

In order to consider profitability we need to consider three examples of borrower payment behaviour:

1. *Complete payback of the loan*: The borrower pays the Mortgage in Full over the prescribed term of the loan
2. *Default on the loan*: The borrower makes payments for a period of time and stops payment. The Lender receives the collateral for the loan (that is the Property) from the borrower.
3. *Early Repayment of the Loan in Full*: The borrower makes payments until a certain point in time and then pays the full Outstanding Balance in a single payment.

Complete Payback

The most common type of mortgage is for a fixed Principal Amount, with a specified Term and contract Rate. We therefore consider the loan details as follows:

Loan of £100,000 with a Term of 25 years and fixed initial contract rate of 6% APR

The borrower maintains the scheduled payments throughout the lifetime of the loan.

Treating the Mortgage as an Annuity and using the Equation (A.1) the monthly payment will be

$$m = \frac{£100000}{\left[\frac{1-(1+(\frac{0.06}{12}))^{-300}}{(\frac{0.06}{12})}\right]} = £644.30$$

From which again by application of Equation (A.1)

$$PV(Payments) = £644.30\left[\frac{1-(1+(\frac{0.03}{12}))^{-300}}{(\frac{0.03}{12})}\right] = £644.30(210.87) = £135,867.70$$

That is from a Return on Capital ROC perspective:

The Mortgage Loan Profitability is £135,867.70 - £100,000 = £35,867.70 in today's £ or a 35.9% return on the £100,000 capital invested.

Loan Default

When a borrower violates the terms and conditions of the mortgage contract by failing to make the scheduled payments, the borrower is considered to be in default. Default may be due to an inability of the borrower to pay the scheduled payments as a result of for example unemployment, or long term illness.

In addition even a borrower with the ability to pay the scheduled amounts may default if the value of the property is worth less than the outstanding balance on the Mortgage loan (the position of negative equity).

For illustration consider a Mortgage loan of £100,000 secured on a property with a current estimated value of £150,000. As long as the value of the home is greater than the outstanding balance on the Mortgage loan, a borrower who could not make the monthly payments could avoid default by selling the property and paying the outstanding balance.

However, if the value of the property is currently estimated to have a value of £60,000 the payment of the loan over the full term would mean paying more for the property than it is worth and so a borrower in this position may choose to default.

Return On Capital can be seen then to be affected not only by the reduced number of payments but also by the loss incurred because the property sold through default proceedings is worth less than the outstanding balance.

The ROC of the defaulted loan thus requires the PV of the payments and the PV of the loss incurred at default. When the borrower defaults, the property is repossessed by the lender and sold, invariably for less than the outstanding balance. In this case we need to consider ROC to be calculated in the following way:

$$ROC = \frac{\sum PV(Payments\ up\ to\ default) + PV(Collateral\ at\ Default) - Capital}{Capital}$$

$$= \frac{\sum PV(m_1, m_2, \ldots, m_t) + PV(\theta_t) - Capital}{Capital}$$

Where m_1, m_2, ..., m_t represent the monthly payments paid prior to default and θ_t represents the value of the property at default with t being the month of default. Consider then a borrower who defaults after 10 years i.e. t = 120 and suppose the property is valued at £60,000 at default i.e. $\theta_{120} = 60000$

Then by Equation (A.1):

Present value of payments to default =

$$£644.30\left[\frac{1-(1+(\frac{0.03}{12}))^{-120}}{(\frac{0.03}{12})}\right] = £644.30(103.56) = £66,724.84$$

And Present value of Property at default $PV(\theta_{120}) = \dfrac{£60000}{(1+(\frac{0.03}{12}))^{120}} = £44,465.74$

Giving $ROC = \dfrac{66,724.84 - 44,465.74 - 100,000}{100,000} = 0.1119 = 11.19\%$

That is from a Return on Capital ROC perspective:

The Mortgage Loan Profitability is 11.2% return on the £100,000 capital invested.

Indicating as expected, borrowers in default are very costly, as the lender is not receiving the revenues that were agreed to upon origination of the mortgage loan. There is thus an obvious benefit to the lender if they can identify the borrowers who are more likely to default and so not originate mortgage loans to these borrowers.

Early Repayment in Full

A borrower may decide, after a period of time has elapsed, that the conditions set out when their mortgage loan started are not now ideal for them. They may then choose to pay off their current loan with the proceeds from another loan using the same property as security. This is the process by which a borrower may refinance the mortgage on more favourable terms and is called Prepayment. This process is popular with borrowers in a declining interest rate environment.

Occasionally however, a borrower might choose to prepay their current mortgage loan from cash in hand (say by an unexpected financial windfall) rather from the proceeds of another loan.

Consider then a borrower who chooses to pay off their mortgage earlier than anticipated after 10 years i.e. t = 120 months by paying the full outstanding balance in a single payment

For this borrower position the Return On Capital can be calculated as follows:

$$ROC = \frac{\sum PV(Payments\ up\ to\ \mathrm{Pr}epayment) + PV(Outs\tan ding\ Balance) - Capital}{Capital}$$

$$= \frac{\sum PV(m_1, m_2, \ldots, m_t) + PV(OB_t) - Capital}{Capital}$$

Where m_1, m_2, ..., m_t represent the monthly payments paid prior to prepayment and OB_t represents the Outstanding Balance in month of prepayment with t being the month of prepayment.

Now the PV (outstanding balance) will be the equal to PV (Remaining mortgage payments).

From Equation (A.1)

$$PV(remaining\ Payments) = \frac{£644.30\left[\frac{1-(1+(\frac{0.06}{12}))^{-12(25-10)}}{(\frac{0.06}{12})}\right]}{(1+(\frac{0.03}{12}))^{12(10)}} = \frac{£76,351.81}{(1+(\frac{0.03}{12}))^{120}} = £56,584.12$$

Which means in Present Value terms the mortgage can be paid off at the end of year 10 for the amount £56,584.12.

Therefore

$$ROC = \frac{66,724.84 + 56,584.12 - 100000}{100000} = 0.2331 = 23.31\%$$

That is from a Return on Capital ROC perspective:

The Mortgage Loan Profitability is 23.3% return on the £100,000 capital invested.

Indicating as expected, that since the capital is all paid back (earlier than anticipated), the ROC is not as high as the ROC for the complete loan payback situation.

Appendix B: Mechanics of Fixed Rate Mortgages

Table 54: Characteristics of a Fixed-Rate Mortgage at 6 Percent*

Remaining Payment	Total payment (£)	Interest (£)	Principal (£)	principal (£)
1	1,288.61	1,000.00	288.61	199,711.90
2	1,288.61	998.56	290.05	199,421.85
120	1,288.61	766.14	522.47	152,704.34
151	1,288.61	678.77	609.84	135,144.78
163	1,288.61	641.16	647.45	127,584.53
202	1,288.61	502.14	786.47	99,641.51
240	1,288.61	338.03	950.58	66,653.87
300	1,288.61	6.41	1,282.20	0.00
Total	386,583.00	186,582.50	200,000.50	—

NOTE: *Based on 25-year maturity.

Consider the purchase of a house with a total cost of H_0 = £250,000 using a loan with a 20 percent down payment, χ = 0.20; an interest rate of r = 6 percent annually; and a N=25-year maturity. This mortgage loan is for £200,000.50. The tabulation above illustrates the changes in interest and principal payments per month over the length of the mortgage contract. In effect an abbreviated amortization schedule. The first two rows of the Table show the mortgage payment in the first and second months of the contract. The monthly payment on this mortgage is £1,288.61. In the first period, £1000.00 of the monthly payment goes to interest rate payments.

This means the principal payment is only: £288.60. Now, let us consider the mortgage payment 10 years into the mortgage. Although the monthly payment does not change, the principal payment has increased to £522.47 and the interest payment component has decreased to £766.14. After 10 years, the homeowner has paid off only £47,296.16 of the original £200,000.50 loan. The month after the halfway point in the mortgage occurs at period 151. The interest payment component of the monthly payment still exceeds the principal payment.

In payment period 163—that is 13 years and 7 months into the contract—the principal component of the monthly payment finally exceeds the interest payment component. From this point forward, the principal payment will be larger than the interest payment. At the end of 20 years, or period 240, the principal component of the £1,288.61 monthly payment is £950.58.

However, £66,653.87 is still owed on the original £200,000.50 loan. The outstanding loan balance does not drop below £100,000 until payment period 202. With a standard 25-year mortgage contract, it takes nearly 17 years to pay off half the mortgage loan. The remaining half of the mortgage will be repaid in the final 8 years of this mortgage.

Table 55: Characteristics of a Fixed-Rate Mortgage at 7 Percent*

Remaining Payment	**Total payment (£)**	**Interest (£)**	**Principal (£)**	**principal (£)**
1	1,413.56	1,166.67	246.89	199,753.61
2	1,413.56	1,165.23	248.33	199,505.28
120	1,413.56	920.27	493.29	157,267.19
181	1,413.56	710.18	703.38	121,041.55
182	1,413.56	706.08	707.48	120,334.07
209	1,413.56	585.77	827.79	99,590.11
240	1,413.56	422.21	991.35	71,387.61
300	1,413.56	8.20	1,405.36	0.00
Total	424,068.00	224,067.50	200,000.50	—

NOTE: *Based on 25-year maturity.

This table shows the standard 25-year mortgage contract if the mortgage interest rate increases from 6 percent to 7 percent. The table shows that a 1% increase in the interest rate increases the monthly mortgage payment from £1,413.56 to £1,288.61 — an £124.95 increase. Furthermore, the increase in the interest rate results in additional back-loading of principal payments. After 10 years, less than £50,000 of the original balance is paid off. The month after the halfway point in the mortgage occurs at period 181. The payment period when the principal component exceeds the interest component does not occur until period 182.

In fact, the outstanding balance will not drop below £100,000 until payment 209—7 months later than if the interest rate is 6 percent.

At the end of 20 years, or period 240, the principal component of the £1413.56 monthly payment is £991.35. However, £71,387.61 is still owed on the original £200,000.50 loan.

This table clearly illustrates the impact of interest rate changes on a mortgage loan. If the total interest payments on the mortgage contract presented in the preceding two tables are compared the 1 percent increase in the interest rate results in £37,503 of additional mortgage payments over the life of the mortgage.

In the following illustration a balloon contract with a 10-year interest-only loan that is rolled into a 15-year fixed-payment mortgage. The table below shows the payment profiles for this contract. We also assume an interest rate of 6 percent and a 20 percent down payment. The interest-only part of the loan requires 120 mortgage payments of £1,000.00 just to cover the interest obligations on the £200,000.50 loan.

Table 56: Characteristics of a Balloon Mortgage at 6 Percent*

Remaining Payment	Total payment (£)	Interest (£)	Principal (£)	principal (£)
1	1,000.00	1,000.00	0.00	200,000.50
2	1,000.00	1,000.00	0.00	200,000.50
120	1,000.00	1,000.00	0.00	200,000.50
121	1,687.72	1,000.00	678.72	199,312.78
163	1,687.72	839.74	847.98	167,100.49
230	1,687.72	503.29	1,184.43	99,473.11
240	1,687.72	442.72	1,245.00	87,298.15
300	1,687.72	8.40	1,679.32	0.00
Total	322,526.40	122,525.90	200,000.50	—

NOTE: *Based on 25-year maturity, 10 years interest only.

After 10 years, the mortgage payment increases to £1,687.72 because the 10-year balloon loan is rolled into a 15-year FRM. Payment number 163 denotes the month in which principal payments exceed interest payments.

In period 230, half of the £200,000.50 debt will be paid off. With this type of mortgage contract, it takes more than 19 years to accrue £100,000 in equity.

The following illustrates the mechanics of a standard variable rate mortgage SVRM used in recent years have a very short period of interest-only payments. The table below presents the payment profiles for a 3-year interest-only SVRM that rolls into a 22-year standard FRM. The assumptions for the interest rate, total contract length, and down payment remain unchanged.

Table 57: Characteristics of an Adjustable-Rate Mortgage with a Constant Interest Rate of 6 Percent*

Remaining Payment	**Total payment (£)**	**Interest (£)**	**Principal (£)**	**principal (£)**
1	1,000.00	1,000.00	0.00	200,000.50
2	1,000.00	1,000.00	0.00	200,000.50
36	1,000.00	1,000.00	0.00	200,000.50
37	1,366.15	998.18	366.17	199,634.35
120	1,366.15	812.24	553.91	161,893.85
151	1,366.15	719.62	646.53	143,277.58
163	1,366.15	679.74	686.41	135,262.37
209	1,366.15	502.73	863.42	99,683.31
240	1,366.15	358.37	1,002.78	71,672.78
300	1,366.15	6.80	1,359.35	0.00
Total	396,663. 60	196,663.10	200,000.50	—

NOTE: *Based on 25-year maturity, 3 years interest only.

The monthly interest payments for this interest-only SVRM are £1,000.00.

Once the standard 22-year mortgage contract takes effect the monthly mortgage payment increases by £366.15 to £1,366.15. This increase is not caused by an interest rate increase, but rather payment toward principal.

In the final illustration below mortgage interest rates have begun to increase recently. What effect does this have on an interest-only SVRM? To show this effect, we allow the interest rate to increase to 7 percent for the standard FRM that is obtained after the 3-year SVRM expires. Table below presents the various payment patterns.

Table 58: Characteristics of an Adjustable-Rate Mortgage with a Rising Interest Rate*

Remaining Payment	Total payment (£)	Interest (£)	Principal (£)	principal (£)
1	1,000.00	1,000.00	0.00	200,000.50
2	1,000.00	1,000.00	0.00	200,000.50
36	1,000.00	1,000.00	0.00	200,000.50
37	1,486.85	1,166.67	320.18	199,680.32
120	1,486.85	967.98	518.87	165,421.13
151	1,486.85	865.46	621.39	147,743.55
182	1,486.85	742.68	744.17	126,573.12
215	1,486.85	585.22	901.63	99,421.66
240	1,486.85	444.10	1,036.70	76,131.74
300	1,486.85	8.62	1,478.23	0.00
Total	428,528.40	228,527.90	200,000.50	—

NOTE: *Based on 25-year maturity, 3 years interest only at a 6 percent interest rate, and the remaining years at 7 percent.

A 100-basis-point increase in the interest rate causes the monthly payment to increase to £1,486.85 from £1,366.15 — a 48 percent increase in the mortgage payment from the interest-only payments. This example illustrates the risk facing homeowners when the interest rate increases before the transition to a standard FRM.

Appendix C: Financial Options and Terminology

Options, which are agreements to buy or sell at a specified future time a particular asset at a specified price, are termed *Forward Contracts*.

The buyer in the contract is said to be in a *long position*, they have a positive amount of the asset or a positive exposure. The seller is said to hold a *short position*, they have a negative exposure.

There are two main types of options

i) *A Call Option*, which gives the holder the right but not the obligation to buy a particular asset for a given amount, the exercise or strike price, at a specified time in the future, the expiration date and

ii) *A Put Option*, which gives the holder the right but not the obligation to sell a particular asset for a given amount, at a specified time in the future.

Options that can only be exercised only on a specified expiry date are referred to as *European* Options, options that can be exercised on and at any time prior to the exercise date are referred to as *American* Options.

Define K to be the strike or exercise price and T the expiration date, then an option's payoff or value V_T at expiration date is:

$$V_T = \begin{cases} \max(S_T - K, 0) \; foraCallOption \\ \max(K - S_T, 0) \; foraPutOption \end{cases}$$

where S_T denotes the price of the underlying asset at maturity t = T.

For the Call Option: the call option will be exercised if the price of the stock is greater than the exercise price, that is, $S_T > K$.

If the stock price was lower than the exercise price then the option would not be exercised and would expire worthless and so $V_T = \max(S_T - K, 0)$. As the stock price rises the option will become more 'in-the- money' and the greater will be the payoff.

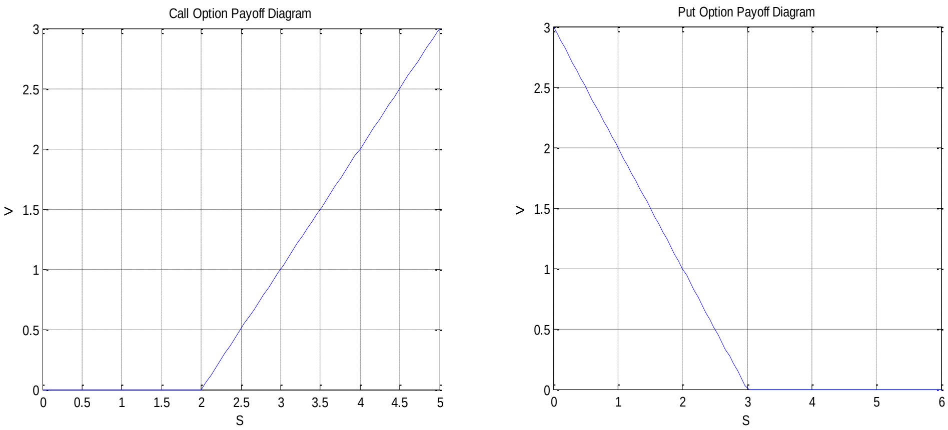


Figure 39: The Payoff diagrams for Call and Put Options

The Payoff diagrams for Call and Put Options are illustrated above are illustrated in Figure 39. above.

Thus far this discussion has focussed on the value of the Call or Put option at expiry but what is of interest is the value of the option now, at a date prior to the expiry date. This value will depend on both the value of the underlying asset today and the time until expiry.

As we have previously seen the value of a Call or Put option at expiry given by

$$V(S_T, T) = V_T = \begin{cases} \max(S_T - K, 0) \ foraCallOption \\ \max(K - S_T, 0) \ foraPutOption \end{cases}$$

where the notation is changed slightly to show that S and t are the variables that affect the value of the option.

However, there are also other parameters that affect the options value, namely, the interest rate and asset price volatility. The interest rate affects the option through the time value of money and the volatility is a measure of how much the underlying asset fluctuates giving essentially a measure of the element of randomness in the asset.

Appendix D: Derivation of the Asset Valuation PDE

This Appendix demonstrates the derivation of the asset valuation PDE using standard no arbitrage arguments. The PDE for the valuation of a mortgage, whose value is a function of interest rate *r*, House price *H* and time *t* can be found as follows:

Let *V(H, r, t)* be the value of the Mortgage option contract at time *t*.

Let Π be the portfolio:

Long one mortgage option position *V* and a short position in each of

(i) Some quantity Δ of the house asset *H*

(ii) Some quantity Δ_1 of another option on the interest rate *r*, with value $V_1(H, r, t)$

Then $\Pi = V - \Delta H - \Delta_1 V_1$

The change in this portfolio $d\Pi$ in a time *dt* is given by changes that are due in part to the option values $dV \textit{ and } \Delta_1 dV_1$ as well as changes in the house value ΔdH and the rental flow received $\Delta\phi H dt$ thus:

$$d\Pi = dV - \Delta dH - \Delta\phi H dt - \Delta_1 V_1$$

Now using $It\hat{o}'s$ Lemma for the two stochastic variables given previously (see Ito (1951) for details):

$$d\Pi = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial H}dH + \frac{\partial V}{\partial r}dr + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2}dt + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V}{\partial r^2}dt + \rho\sigma_H\sigma_r H\sqrt{r}\,\frac{\partial^2 V}{\partial H\,\partial r}dt$$

$$-\Delta dH - \Delta\phi H dt$$

$$-\Delta_1$$

$$\left(\frac{\partial V_1}{\partial t}dt + \frac{\partial V_1}{\partial H}dH + \frac{\partial V_1}{\partial r}dr + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V_1}{\partial H^2}dt + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V_1}{\partial r^2}dt + \rho\sigma_H\sigma_r H\sqrt{r}\,\frac{\partial^2 V_1}{\partial H\,\partial r}dt\right)$$

Collecting together common deterministic terms (dt) and common stochastic terms (dH) and (dr) we get:

$$d\Pi = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V}{\partial H\,\partial r} - \Delta\phi H)dt$$

$$-\Delta_1 (\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V_1}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V_1}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V_1}{\partial H\,\partial r})dt$$

$$+ (\frac{\partial V}{\partial H} - \Delta - \Delta_1 \frac{\partial V_1}{\partial H})dH + (\frac{\partial V}{\partial r} - \Delta_1 \frac{\partial V_1}{\partial r})dr$$

We now seek to eliminate the uncertainty (risk) in the portfolio by choosing:

$$(\frac{\partial V}{\partial r} - \Delta_1 \frac{\partial V_1}{\partial r}) = 0 \text{ giving } \Delta_1 = (\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r})$$

and given this Δ_1 by choosing

$$(\frac{\partial V}{\partial H} - \Delta - \Delta_1 \frac{\partial V_1}{\partial H}) = 0 \text{ so that } \Delta = [\frac{\partial V}{\partial H} - (\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r})\frac{\partial V_1}{\partial H}]$$

On back substitution of these expressions we obtain a risk free portfolio, which by the No Arbitrage principle[19] must have a value equal to that of a risk free interest bearing account with $d\Pi = r\Pi dt = r(V - \Delta H - \Delta_1 V_1)dt$

Therefore:

$$d\Pi = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V}{\partial H\,\partial r})dt$$

$$-[\frac{\partial V}{\partial H} - (\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r})\frac{\partial V_1}{\partial H}]\phi H dt$$

$$-(\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r}) (\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V_1}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V_1}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V_1}{\partial H\,\partial r})dt$$

[19] No arbitrage arguments implies that the return on the portfolio to be $r\Pi dt$ since the growth of the portfolio in a time step dt is equal to the risk-free growth rate of the portfolio, as the portfolio is now completely deterministic (volatility term is zero).

$$= (rV - [\frac{\partial V}{\partial H} - (\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r})\frac{\partial V_1}{\partial H}]rH - (\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r})rV_1)dt$$

Now dividing throughout by *dt* and collecting together like terms gives:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V}{\partial H\,\partial r} + (r-\phi)H\frac{\partial V}{\partial H} - rV$$

$$= (\frac{\partial V}{\partial r} \ / \ \frac{\partial V_1}{\partial r})$$

$$(\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V_1}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V_1}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V_1}{\partial H\,\partial r} + (r-\phi)H\frac{\partial V_1}{\partial H} - rV_1)$$

Dividing throughout by $(\frac{\partial V}{\partial r})$ gives:

$$(\frac{\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V}{\partial H\,\partial r} + (r-\phi)H\frac{\partial V}{\partial H} - rV}{\frac{\partial V}{\partial r}})$$

$$=$$

$$(\frac{\frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V_1}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V_1}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\frac{\partial^2 V_1}{\partial H\,\partial r} + (r-\phi)H\frac{\partial V_1}{\partial H} - rV_1}{\frac{\partial V_1}{\partial r}})$$

Remembering that *V* and V_1 are the prices of 2 arbitrary derivatives the LHS is a function of *V* and not V_1 and the RHS is a function of V_1 and not *V*.

The only way that this ratio is constant for all derivative pairs is if it does not actually depend on them, that is we are independent of the contract type (maturity date). That is, it is a constant function *a(H, r, t)* which is interpreted as being given the form

$$a(H,r,t) = \sigma_r \sqrt{r}\lambda(r,t) - \gamma(\theta - r)$$

where $\lambda(r,t)$ is termed the Market Price of interest-rate risk.

Now the effect of the local expectations hypothesis (LEH) is that we are assured that this interest rate factor disappears, that is $\lambda(r,t) = 0$ giving $a(H,r,t) = -\gamma(\theta - r)$

Thus, dropping subscripts we have that

$$\frac{\dfrac{\partial V}{\partial t} + \dfrac{1}{2}\sigma_H^2 H^2 \dfrac{\partial^2 V}{\partial H^2} + \dfrac{1}{2}\sigma_r^2 r \dfrac{\partial^2 V}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\,\dfrac{\partial^2 V}{\partial H\,\partial r} + (r-\phi)H\dfrac{\partial V}{\partial H} - rV}{\dfrac{\partial V}{\partial r}}$$

$$= a(H,r,t) = -\gamma(\theta - r)$$

hence:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_H^2 H^2 \frac{\partial^2 V}{\partial H^2} + \frac{1}{2}\sigma_r^2 r \frac{\partial^2 V}{\partial r^2} + \rho\sigma_H \sigma_r H\sqrt{r}\,\frac{\partial^2 V}{\partial H\,\partial r} + (r-\phi)H\frac{\partial V}{\partial H} - rV$$

$$= -\gamma(\theta - r)\frac{\partial V}{\partial r}$$

So that we arrive finally at the PDE as given previously

$$\frac{1}{2}H^2\sigma_H^2\frac{\partial^2 V}{\partial H^2} + \rho H\sqrt{r}\sigma_H\sigma_r\frac{\partial^2 V}{\partial H\,\partial r} + \frac{1}{2}r\sigma_r^2\frac{\partial^2 V}{\partial r^2} + \gamma(\theta - r)\frac{\partial V}{\partial r} + (r-\phi)H\frac{\partial V}{\partial H} + \frac{\partial V}{\partial t} - rV = 0$$

This PDE will be solved using a finite difference methodology stepping backwards in time, where it is necessary to start the process from the known information at maturity, referring to the known cash flow at the final moment of the contract.

Appendix E: Taylor Series Expansion

The Taylor Series Expansion of a function $f(x)$ about a point in space x_j is given by

$$f(x_j + \Delta x) = \sum_{n=0}^{\infty} \frac{\Delta x^n}{n!} \frac{\partial^n f}{\partial x^n}\bigg|_{x=x_j}$$

$$= f(x_j) + \Delta x \frac{\partial f}{\partial x}\bigg|_{x=x_j} + \frac{\Delta x^2}{2} \frac{\partial^2 f}{\partial x^2}\bigg|_{x=x_j} + \frac{\Delta x^3}{6} \frac{\partial^3 f}{\partial x^3}\bigg|_{x=x_j} + \quad ...$$

If $x_{j+1} = x_j + \Delta x$ then we can write the above equation as

$$f(x_{j+1}) = f(x_j) + \Delta x \frac{\partial f}{\partial x}\bigg|_{x=x_j} + \frac{\Delta x^2}{2} \frac{\partial^2 f}{\partial x^2}\bigg|_{x=x_j} + \quad HigherOrderTerms$$

Rearranging and letting $f(x_j) = f_j$ and $f(x_{j+1}) = f_{j+1}$ the Taylor series expansion may be written

$$\frac{\partial f}{\partial x}\bigg|_j = \frac{f_{j+1} - f_j}{\Delta x} - \frac{\Delta x}{2} \frac{\partial^2 f}{\partial x^2}\bigg|_j + HigherOrderTerms$$

Now using the standard δ and 'Big O' notation the approximation can be represented as

$\frac{\delta f}{\delta x}\bigg|_j = \frac{f_{j+1} - f_j}{\Delta x}$. The truncation error is $-\frac{\Delta x}{2} \frac{\partial^2 f}{\partial x^2}\bigg|_j + HigherOrderTerms$

and the leading term in the truncation error is given by $-\frac{\Delta x}{2} \frac{\partial^2 f}{\partial x^2}\bigg|_j$ so that we may now write $\frac{\partial f}{\partial x}\bigg|_j = \frac{f_{j+1} - f_j}{\Delta x} + O(\Delta x)$

This is now interpreted as saying that $\left.\frac{\delta f}{\delta x}\right|_j = \frac{f_{j+1} - f_j}{\Delta x}$ is a first order approximation to the first derivative at $x = x_j$, since the error is proportional to Δx.

The following approximations are valid for the derivative of a function *f(x, t)*.

A *forward difference* approximation of a derivative is the approximation

$\left.\frac{\partial f}{\partial x}\right|_j = \frac{f_{j+1} - f_j}{\Delta x} + O(\Delta x)$ derived from

$$f(x_j + \Delta x) = \sum_{n=0}^{\infty} \frac{\Delta x^n}{n!} \left.\frac{\partial^n f}{\partial x^n}\right|_{x=x_j}$$

A *backward difference* approximation of a derivative is the approximation

$\left.\frac{\partial f}{\partial x}\right|_j = \frac{f_j - f_{j-1}}{\Delta x} + O(\Delta x)$ derived from

$$f(x_j - \Delta x) = \sum_{n=0}^{\infty} (-1)^n \frac{\Delta x^n}{n!} \left.\frac{\partial^n f}{\partial x^n}\right|_{x=x_j}$$

And a *central difference* approximation of a derivative is the approximation

$\left.\frac{\partial f}{\partial x}\right|_j = \frac{f_{j+1} - f_{j-1}}{2\Delta x} + O(\Delta x^2)$ derived from taking the average of the forward and backward difference formulae yielding

$$2\left.\frac{\partial f}{\partial x}\right|_j = \frac{f_{j+1} - f_j}{\Delta x} - \frac{\Delta x}{2} \left.\frac{\partial^2 f}{\partial x^2}\right|_j + \frac{f_j - f_{j-1}}{\Delta x} + \frac{\Delta x}{2} \left.\frac{\partial^2 f}{\partial x^2}\right|_j + O(\Delta x^2)\,.$$

The difference equations are obtained by removing the error terms indicated by the '*Big O*' notation. The order of the error for each of the equations is easily seen by considering the Taylor series expansions given above.

The terms $O(\Delta x)$ and $O(\Delta x^2)$ indicate the truncation error of the difference equations. Both forward and backward difference approximations have an error $O(\Delta x)$ and the central difference approximation has error $O(\Delta x^2)$. If a better approximation is required, the computational mesh (on which the approximated solution is calculated), can be made finer (by making Δx smaller) or information can be added by including higher-order neighbouring terms, which will involve additional mesh points.

In order to simplify further the notation, when convenient, the discretisation points will be labelled with appropriate indices so that for example with $f_j^n \equiv f(x_j, t_n)$ where $x_j = j\Delta x$ and $t_n = n\Delta t$ the central difference approximation becomes $\frac{\partial f}{\partial x} \approx \frac{f_{j+1}^n - f_{j-1}^n}{2\Delta x}$.

This approach is readily extended to obtain similar approximations for second derivatives i.e. $\frac{\partial^2 f}{\partial x^2} \approx \frac{f_{j+1}^n - 2f_j^n + f_{j-1}^n}{\Delta x^2}$ and can be extended even further to encompass functions of more than a single variable.

For example with $f_{j,k}^{n} \equiv f(x_j, y_k, t_n)$ where $x_j = j\Delta x$; $y_k = k\Delta y$ and $t_n = n\Delta t$ the central difference approximation becomes $\dfrac{\partial f}{\partial x} \approx \dfrac{f_{j+1,k}^{n} - f_{j-1,k}^{n}}{2\Delta x}$ and $\dfrac{\partial f}{\partial y} \approx \dfrac{f_{j,k+1}^{n} - f_{j,k-1}^{n}}{2\Delta y}$ for the first derivatives.

In the time dimension we would have $\dfrac{\partial f}{\partial t} \approx \dfrac{f_{j,k}^{n+1} - f_{j,k}^{n}}{\Delta t}$.

Appendix F: Time Series Data and Credit Score Information

Table 59: Time Series Data for the Financial Times House Price Index

Month	%Change FT_HPI	Discount Factor
Jan-02	1.4	1
Feb-02	1.4	1
Mar-02	1.6	1
Apr-02	1.7	1
May-02	1.9	1
Jun-02	1.9	1
Jul-02	2	1
Aug-02	1.8	1
Sep-02	1.9	1
Oct-02	1.7	1
Nov-02	1.5	1
Dec-02	1.6	1

Source: Academetrics Bulletins

Table 60: Time Series Data for the UK leading Economic Indicator

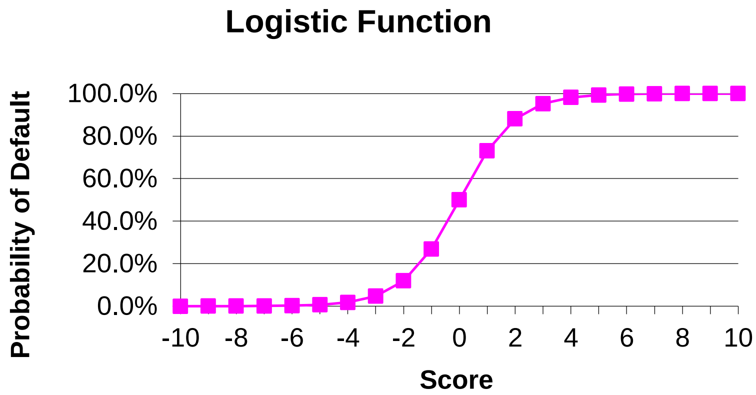
Month	%Change LEI	Discount Factor
Jan-02	-0.2	0.998
Feb-02	-0.2	0.998
Mar-02	-0.3	0.997
Apr-02	0.6	1
May-02	0.4	1
Jun-02	-0.2	0.998
Jul-02	-0.7	0.993
Aug-02	0.4	1
Sep-02	-0.1	0.999
Oct-02	-0.2	0.998
Nov-02	-0.4	0.996
Dec-02	-0.7	0.993

Source: Conference Board Bulletins

Table 61: From Application Score to Probability of Default using simple Logistic Transformation

Quality	Internal Rating	Re-Scaled	Logistic PD
Low	100	0.0	50.00%
	200	-1.0	26.89%
	300	-2.0	11.92%
	400	-3.0	4.74%
	500	-4.0	1.80%
	600	-5.0	0.67%
	700	-6.0	0.25%
	800	-7.0	0.09%
	900	-8.0	0.03%
High	1000	-9.0	0.01%

Figure 40: Default Probability associated with Application Score



Appendix G: Derivation of Affordability Model Formulae

(1) For correlated z_L and z_E:

The random variables z_L and z_E are correlated with correlation coefficient ρ.

Under our Gaussian assumptions, the probability of default is described by a Bivariate Normal Distribution derived as follows:

$z_L = \frac{x_L - \mu_L}{\sigma_L}$ and $z_E = \frac{x_E - \mu_E}{\sigma_E}$ are one to one and have inverse

$x_L = \mu_L + \sigma_L z_L$ and $x_E = \mu_E + \sigma_E z_E$ with $-\infty < x_L, x_E < +\infty$ and $-\infty < z_L, z_E < +\infty$

The Jacobian $J(x_L, x_E; z_L, z_E) = \begin{bmatrix} \frac{\partial x_L}{\partial z_L} & \frac{\partial x_L}{\partial z_E} \\ \frac{\partial x_E}{\partial z_L} & \frac{\partial x_E}{\partial z_E} \end{bmatrix} = \begin{bmatrix} \sigma_L & 0 \\ 0 & \sigma_E \end{bmatrix}$

From which $|J(x_L, x_E; z_L, z_E)| = \sigma_L \sigma_E$

Now from the Joint p. d. f . of x_L and x_E we have

$$f_{x_L, x_E}(x_L, x_E) =$$

$$\frac{1}{2\pi\sigma_L\sigma_E\sqrt{1-\rho^2}} \exp\left\{\left[\frac{-1}{2(1-\rho^2)}\right]\left[\left(\frac{x_L-\mu_L}{\sigma_L}\right)^2 - 2\rho\left(\frac{x_L-\mu_L}{\sigma_L}\right)\left(\frac{x_E-\mu_E}{\sigma_E}\right) + \left(\frac{x_E-\mu_E}{\sigma_E}\right)^2\right]\right\}$$

then the probability of default is

$$f_{z_L, z_E}(z_L, z_E) = |J(x_L, x_E; z_L, z_E)|\, f_{x_L x_E}(\mu_L + \sigma_L z_L, \mu_E + \sigma_E z_E)$$

Substituting and rearranging gives the probability of default at time t as

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}Q\right] \text{ where } Q = \frac{1}{1-\rho^2}\left[z_L^2 - 2\rho z_L z_E + z_E^2\right]$$

(2) Expected Loss

Let $g(x_E)$ be the probability density of x_E.

Then for any K: $V_{BG} = V_{LL} = E[\max(0, K - x_E)] = \int_{-\infty}^{K} (K - x_E) g(x_E) dx_E$

Now by assumption ln(x_E)~N $(mt,\sigma_E^2 t)$ and by the properties of the lognormal distribution the mean of *ln(x_E)* is *mt* where $mt = \ln(E(x_E)) - \frac{\sigma_E^2 t}{2}$ with the mean of x_E being $\exp\left(mt + \sigma_E^2 t \big/ 2\right)$

Define the standard normal variable

$Q = \frac{\ln(x_E) - mt}{\sigma_E \sqrt{t}}$ with density $h(Q) = \frac{1}{\sqrt{2\pi}} \exp\left(-Q^2 \big/ 2\right)$

Then using the change of variables we can write

$$E[\max(0, K - x_E)] = \int_{-\infty}^{\frac{\ln k - mt}{\sigma_E \sqrt{t}}} (K - e^{Q\sigma_E \sqrt{t} + mt}) h(Q) dQ$$

$$= K \int_{-\infty}^{\frac{\ln K - mt}{\sigma_E \sqrt{t}}} h(Q) dQ - \int_{-\infty}^{\frac{\ln K - mt}{\sigma_E \sqrt{t}}} e^{Q\sigma_E \sqrt{t} + mt} h(Q) dQ$$

But

$$\exp(Q\sigma_E \sqrt{t} + mt) h(Q) = \frac{1}{\sqrt{2\pi}} \exp\left(\tfrac{1}{2}\left(-Q^2 + 2Q\sigma_E \sqrt{t} + 2mt\right)\right)$$

$$= \frac{1}{\sqrt{2\pi}} \exp \frac{1}{2}\left(-\left(Q - \sigma_E \sqrt{t}\right)^2 + 2mt + \sigma_E^2 t\right)$$

$$= \frac{\exp\left(mt + \frac{1}{2}\sigma_E^2 t\right)}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}(Q - \sigma_E \sqrt{t})^2\right)$$

$$= \exp\left(mt + \tfrac{1}{2}\sigma_E^2 t\right) h(Q - \sigma_E \sqrt{t})$$

Therefore

$$E[\max(0, K - x_E)]$$

$$= K \int_{-\infty}^{\frac{\ln K - mt}{\sigma_E \sqrt{t}}} h(Q) dQ - \exp(mt + \tfrac{1}{2}\sigma_E^2 t) \int_{-\infty}^{\frac{\ln K - mt}{\sigma_E \sqrt{t}}} h(Q - \sigma_E \sqrt{t}) dQ$$

$$= K\left[\Phi\left(\frac{\ln K - mt}{\sigma_E \sqrt{t}}\right)\right] - \exp\left(mt + \tfrac{1}{2}\sigma_E^2 t\right)\Phi\left(\frac{\ln K - mt}{\sigma_E \sqrt{t}} - \sigma_E \sqrt{t}\right)$$

$$= K\Phi\left(-\left(\frac{-\ln K + \ln(E(x_E)) - \frac{1}{2}\sigma_E^2 t}{\sigma_E \sqrt{t}}\right)\right) - E(x_v)\Phi\left(-\left(\frac{-\ln K + \ln(E(x_E)) - \frac{1}{2}\sigma_E^2 t + \sigma_E^2 t}{\sigma_E \sqrt{t}}\right)\right)$$

$$= K\Phi\left(-\left(\frac{\ln\left(\frac{E(x_E)}{K}\right) - \frac{1}{2}\sigma_E^2 t}{\sigma_E \sqrt{t}}\right)\right) - E(x_E)\Phi\left(-\left(\frac{\ln\left(\frac{E(x_E)}{K}\right) + \frac{1}{2}\sigma_E^2 t}{\sigma_E \sqrt{t}}\right)\right)$$

$$= \Phi\left(-z_E\right) - x_E(t = 0) e^{\mu_E t} \Phi(-z_E')$$

From the Pay-Off we know K=1. From the relationships between the normal and log normal variables we know $E(x_E) = x_E(t=0)\exp(\mu_E t)$

and so $z'_E = z_E + \sigma_E\sqrt{t}$ with $z_E = \frac{\ln(x_E \exp(\mu_E t)) - \frac{1}{2}\sigma_E^2 t}{\sigma_E\sqrt{t}} = \frac{\ln(x_E) + (\mu_E - \frac{1}{2}\sigma_E^2)t}{\sigma_E\sqrt{t}}$

(3) Loss Given Default

$V_{LL} = V_{LGD}\Phi(-z_E)$ and result follows.

Appendix H: Lattice Based Simplistic Default

For the purposes of this illustration assume the following:

- A house has a current value of £100.
- The standard deviation of the return to housing is 0.22314355
- The risk-free interest rate is 4% per annum.
- In order to purchase the house, we promise to repay a lender £95 in 2 years.

Given the previous assumptions, we assume that the house value will either rise to £125 or fall to £80 by the end of the first year (with equal probability).

$$125 = 100 * e^{\sigma\sqrt{\Delta t}} \qquad 80 = 100 * e^{-\sigma\sqrt{\Delta t}}$$

By the end of the second year, the value of the house will be £156.25, £100, or £64.

Average growth rate = 2.5% per year

50% chance of 25% appreciation

50% chance of 20% loss

Table 62: House Price Paths

Year 0	Year 1	Year 2
		£156.25
	£125.00	
£100.00		£100.00
	£80.00	
		£64.00

Binomial Model**:** Cox, Ross, and Rubinstein, 1979 (CRR) – (discrete time version)

$$E[H] = pHu + (1-p)Hd \qquad p = \frac{a-d}{u-d},\ u = e^{\sigma\sqrt{\Delta t}},\ d = e^{-\sigma\sqrt{\Delta t}}, a = (1+r)^{\Delta t}$$

Default Values

At end of year 2, we owe £95 to lender.

- If house value = £156.25, then our equity is £61.25 and we should repay the loan (not default). (£156.25 - £95 = £61.25)
- If house value = £64.00, then our equity is £-31.00 and we should default (lender gets to keep house). (£64.00 - £95 = £-31)

D = min [K,H]

Mortgage Value

Starting with the terminal payoffs, we need to calculate the present value of the mortgage.

Thus, we need to calculate the *pseudo-probability* of change in house prices.

$$p = \frac{a-d}{u-d} = \frac{1.04-0.80}{1.25-0.80} = 0.533333 \qquad 1-p = \frac{u-a}{u-d} = \frac{1.25-1.04}{1.25-0.80} = 0.466666$$

At the end of year 1, the present values of the terminal pay-offs are calculated as:

$$D_u = [pD_{uu} + (1-p)D_{ud}]/(1+r) \qquad D_d = [pD_{du} + (1-p)D_{dd}]/(1+r)$$

Finally, at mortgage origination, the present value of the loan is calculated as:

$$D = [pD_u + (1-p)D_d]/(1+r)$$

Table 63: Present Values

Year 0	Year 1	Year 2
		£95.00
	£91.35	
£81.59		£95.00
	£77.44	
		£64.00

Note: Based on our assumptions of changes in house prices, the lender will originate a mortgage of £81.59 at Year 0.

We borrower £81.59 and promise to repay £95 at the end of Year 2, what is our effective interest rate on this mortgage?

Interest Rate Answer:

$$\text{£}81.59 = \text{£}95/(1+r)^2$$
$$r = 7.9054\%$$

Note: since the risk-free rate is 4% this implies that the default risk premium for this mortgage is 3.9054%

$$r_c = r_{rf} + r_d$$

Remember – risk free rate is 4%.

Thus,

$$£87.83 = £95/(1+0.04)^2$$

What is the default premium to risky borrowers?

Default Premium = Risk-free Value – Risky Value

Premium = £87.83 - £81.59

Premium = £6.24

In other words, this is the value of the default option to the borrower.

Appendix I: Classification Table Calculations

The model classifies an observation as an event if its estimated probability is greater than or equal to a given probability cut points.

Table 64: Output from PROC LOGISTIC; CTABLE with PPROB = (0 to 1 by 0.1)

	Correct		Incorrect		Percentages (%)				
Prob. Level	Event	Non Event	Event	Non Event	Correct	Sensitivity	Specificity	FALSE POS	FALSE NEG
0	57	0	43	0	57	100	0	43	.
0.1	57	1	42	0	58	100	2.3	42.4	0
0.2	55	7	36	2	62	96.5	16.3	39.6	22.2
0.3	51	19	24	6	70	89.5	44.2	32	24
0.4	50	25	18	7	75	87.7	58.1	26.5	21.9
0.5	45	27	16	12	72	78.9	62.8	26.2	30.8
0.6	41	32	11	16	73	71.9	74.4	21.2	33.3
0.7	32	36	7	25	68	56.1	83.7	17.9	41
0.8	24	39	4	33	63	42.1	90.7	14.3	45.8
0.9	6	42	1	51	48	10.5	97.7	14.3	54.8
1	0	43	0	57	43	0	100	.	57
					Tot Correct / Total	*Correct Event/ Tot Event*	*Correct N.Event/ Tot N.Event*	*F.Pos / (F.Pos+Pos)*	*F.Neg / (F.Neg+Neg)*
Item	*a*	*b*	*c*	*d*	*(a+b) / (a+b+c+d)*	*a / (a+d)*	*b / (b+c)*	*c / (a+c)*	*d / (b+d)*

Suppose the cut point is 0.3 so that

a=51 b=19 c=24 d=6 then calculation runs as illustrated

(51+19)/ (51+19+24+6)	51/ (51+6)	19/ (19+24)	24 / (51+24)	6 / (19+6)
70/100	51/57	19/43	24/75	6/25
70%	89.5%	44.2%	32%	24%