

# Mortgage Default during the U.S. Mortgage Crisis\*

Thomas Schelkle†

February 15, 2012

## Abstract

This paper asks which theories of mortgage default are quantitatively consistent with observations in the United States during 2002-2010. Theoretical models are simulated for the observed time-series of aggregate house prices. Their predictions are then compared to actual default rates on prime fixed-rate mortgages. An out-of-sample test discriminates between estimated reduced forms of the two most prominent theories. The test reveals that the double-trigger hypothesis attributing mortgage default to the joint occurrence of negative equity and a life event like unemployment outperforms a frictionless option-theoretic default model. Based on this finding a structural partial-equilibrium model with liquidity constraints and idiosyncratic unemployment shocks is presented to provide micro-foundations for the double-trigger hypothesis. In this model borrowers with negative equity are more likely to default when they are unemployed and have low liquid wealth. The model explains most of the observed strong rise in mortgage default rates. A policy implication of the model is that subsidizing homeowners can mitigate a mortgage crisis at a lower cost than bailing out lenders.

*JEL codes: E21, G21, D11*

*Keywords: Mortgage default, mortgage crisis, house prices, negative equity*

---

\*I thank Francesco Caselli and Rachel Ngai for helpful discussions and advice. I am also grateful for comments received from Wouter den Haan, Albert Marcet, Stephan Seiler and seminar participants at Aachen, Cologne, LSE, Manchester, Royal Holloway and the Econometric Society European Meeting 2011 in Oslo. Special thanks go to Jason Friend for help with the data. All remaining errors are my own. A previous version of this paper was circulated under the title “Simulating Mortgage Default”.

†Economics Department; London School of Economics; Houghton Street; London WC2A 2AE; United Kingdom; Email: t.schelkle@lse.ac.uk.

# 1 Introduction

After the collapse of the house price boom in the United States residential mortgage delinquencies of both prime and subprime loans have increased substantially. The widespread rise in default rates and resulting losses of mortgage-backed-securities marked the onset of the recent financial and economic crisis. These events highlight key research questions on mortgage default. What are the economic mechanisms driving mortgage default? And what explains the strong rise in mortgage default rates in recent years?

This paper examines how well theoretical models of mortgage default can quantitatively explain the rise in default rates in the United States between 2002 and 2010. Theoretical models are simulated for the observed time-series of aggregate house prices and a realistic microeconomic house price distribution. Their predictions are then compared to data on default rates of prime fixed-rate mortgages. In the first part of the paper the observed variation in default rates and aggregate house prices is used to discriminate between the two major mortgage default theories - the frictionless option-theoretic default model and the “double-trigger” hypothesis.

The traditional frictionless option-theoretic literature, sometimes also called the “ruthless” default model, assumes that borrowers default on their mortgage in order to maximize their financial wealth. In this framework negative equity is a necessary, but not sufficient, condition for default. Instead there exists a threshold level of negative equity or the house price such that a rational wealth-maximizing agent will exercise the default option as in Kau, Keenan, and Kim (1994), among others. This theory assumes that the borrower has access to a perfect credit market for unsecured credit such that default is unaffected by liquidity considerations and income fluctuations. Quercia and Stegman (1992) and Vandell (1995) provide a survey and further references.

Another prominent idea on mortgage default is the double-trigger hypothesis. This theory also views negative equity as a necessary condition for default. But it attributes default to the joint occurrence of negative equity and a life event like unemployment or divorce. The double-trigger hypothesis is well-known among mortgage researchers. But it is usually discussed in words or stylized two-period models as in Gerardi, Shapiro, and Willen (2007), Foote, Gerardi, and Willen (2008) and Foote, Gerardi, Goette, and Willen (2009), among others, and has not been presented as a structural dynamic stochastic model.

These two microeconomic theories are tested on their aggregate predictions. The procedure specifies reduced form models of the two theories, estimates them on part of the

data and then tests the estimated models on out-of-sample predictions. The result of this test is that the double-trigger hypothesis outperforms the frictionless default model. The frictionless theory is excessively sensitive to changes in aggregate house prices and predicts a far too strong rise in default rates. In contrast, the double trigger hypothesis is consistent with the evidence. The economic reason is that default rates have increased roughly in proportion to the number of borrowers who experience any level of negative equity as predicted by the double-trigger theory. In contrast, the predictions of the frictionless theory are based on the number of homeowners experiencing extreme levels of negative equity and this has increased by much more than actual default rates. This is an important result in itself given the disagreement in the literature. It is also an important step towards developing mortgage default models that can be used for policy and risk analysis because such analysis needs to be based on models that are empirically accurate.

Based on this finding the second part of the paper aims at providing a micro-foundation for the double-trigger hypothesis. A structural dynamic stochastic partial-equilibrium model of mortgage default featuring liquidity constraints and idiosyncratic unemployment shocks is presented. The liquidity constraint forces unemployed borrowers who have exhausted their buffer stock savings to make painful cuts to consumption. This magnifies the cost of servicing the mortgage such that unemployment becomes a trigger event for default. In addition the model includes a direct utility flow of owning a house. This is an important feature to generate double-trigger behavior because it prevents employed agents from defaulting after a strong fall of house prices. The model then attributes default to the joint occurrence of negative equity and the liquidity problems caused by unemployment. The model is calibrated, estimated and assessed on its power to predict out-of-sample. A comparison to observed default rates reveals that the model can quantitatively explain most of the rise in mortgage default as a consequence of falling aggregate house prices.

One benefit of the structural model is that it can be used for policy analysis. This is exemplified by analyzing two possible policies in a mortgage crisis that neutralize the losses lenders incur from mortgage default. One could either bail out the lenders or mitigate the liquidity problems of homeowners who would otherwise default such that they stay in their houses. An implication of the structural model is that a subsidy policy to homeowners is the cheaper option when liquidity problems play a key role in default decisions.

From a macroeconomic perspective the finding that the structural and reduced-form

double-trigger model can explain the rise in default rates by the dynamics of aggregate house prices is important. This points towards the existence of systematic macroeconomic risk in the mortgage market. The main alternative explanation is that lending standards and loan quality deteriorated sharply before the crisis. This paper presents evidence that at least in my data set on prime mortgages this is an unlikely explanation for the rise in default rates because average loan characteristics are fairly stable over time.

As a background to the paper it is important to know that loan-level data that links an individual borrower's repayment history to the history of individual house prices and employment status does not exist. This makes it difficult to distinguish empirically between different theories at the individual level. This paper takes a different approach and tests the aggregate predictions of different theories. Along this line the key and unique feature of the paper is that it includes a realistic microeconomic house price distribution around the aggregate trend. This means that an empirically successful model is required to be consistent with both the aggregate house price trends and the moments of the observed microeconomic distribution. In contrast, the prior empirical literature relies on regional house price indices as explanatory variables and thus very likely omits part of the microeconomic house price variation from its regressions.

This is the first paper that compares simulations from theoretical mortgage default models directly to empirical observations. The prior literature has in contrast been divided into theoretical work that does not discuss the explanatory power of the theories on the one hand, and reduced-form regressions on the other.

The structural model of the paper builds on previous work by Campbell and Cocco (2003, 2011) and Corradin (2009) who also model liquidity constraints in a mortgage framework.<sup>1</sup> These models are similar to the structural model presented in this paper, but these papers do not compare the models to the data. Their focus is also different, for example Campbell and Cocco (2011) are mainly concerned with theoretical differences between fixed- and variable rate mortgages. In contrast, my paper adds the macroeconomic perspective. It shows how variation in the time-series of aggregate house prices can explain the rise in default rates during recent years within a structural model. My analysis also reveals that in addition to liquidity constraints it is important to allow for a direct utility flow from owning a house as explained above. Otherwise the model remains too close to a ruthless default model and cannot match the data well.

---

<sup>1</sup>In modeling liquidity constraints the structural model also builds on the buffer-stock saving framework of Zeldes (1989), Deaton (1991) and Carroll (1997).

A vast number of empirical papers have studied the determinants of mortgage default typically estimating hazard models on loan-level data. The pre-crisis literature is surveyed by Quercia and Stegman (1992) and Vandell (1995) and an example is the study by Deng, Quigley, and Van Order (2000). The U.S. mortgage crisis has then caused an enormous increase in empirical work on mortgage default.<sup>2</sup> These papers present a wealth of evidence that negative equity or falling house prices are a strong determinant of default. Some studies have also investigated the role of life events as triggers for default and found that state unemployment or divorce rates can explain default. My paper is motivated by these empirical results. But it uses a very different methodology and thus provides complementary evidence on the relative merit of the two theories.

The empirical literature also finds a great heterogeneity in default behavior for borrowers with the same level of negative equity (Quercia and Stegman 1992). The structural model I present here can rationalize this fact because in that model the default threshold of negative equity depends on liquid wealth and employment status. Individual heterogeneity in these variables, which are unobserved in all standard mortgage data sets, then causes heterogeneity in default behavior of borrowers with the same level of negative equity. The theoretical model also suggests that interaction effects between negative equity and variables measuring liquidity are of key importance for default as has been found empirically by Elul, Souleles, Chomsisengphet, Glennon, and Hunt (2010).

The paper is structured as follows. Section 2 describes the data and empirical facts on mortgages and house prices. The test between the two theories based on reduced-form models is presented in section 3. The structural model is developed in section 4 and parameterized in section 5. The results of the structural model are presented in section 6. Section 7 discusses the alternative explanation for the rise in default rates that loan quality deteriorated sharply and shows that there is no strong evidence for this in my data set. The structural model is applied for policy analysis in section 8. For reasons explained in the data section most of the paper concentrates on loans with a high loan-to-value ratio, but section 9 discusses an extension to lower loan-to-value ratios. Section 10 concludes.

---

<sup>2</sup>Studies within this extensive literature differ by research question, estimation method, analyzed data set and results. A detailed literature review that would do justice to these different contributions is unfortunately beyond the scope of this paper. Examples of this empirical research include Amromin and Paulson (2009), Bajari, Chu, and Park (2010), Demyanyk and Van Hemert (2011), Elul, Souleles, Chomsisengphet, Glennon, and Hunt (2010), Foote, Gerardi, Goette, and Willen (2008), Foote, Gerardi, Goette, and Willen (2009), Foote, Gerardi, and Willen (2008), Gerardi, Lehnert, Sherlund, and Willen (2008), Gerardi, Shapiro, and Willen (2007), Ghent and Kudlyak (2010), Guiso, Sapienza, and Zingales (2011), Jagtiani and Lang (2010), Mayer, Pence, and Sherlund (2009) and Mian and Sufi (2009), among others.

## 2 Data and Empirical Facts

This section presents the data on mortgages default rates and house prices and the key facts the paper attempts to explain. It also describes how the simulation procedure for house prices is based on empirical evidence.

### 2.1 Mortgage Data

In this paper, I use aggregate data on mortgage characteristics and payment histories in the United States. The data set contains information that was aggregated from the large loan-level data base of Lender Processing Services (LPS), also known as McDash data. “Aggregate” here simply means that my data contains the average value of a certain characteristic for all loans in the data base that satisfy a set of conditions that I can specify. These conditioning variables allow to select sub-samples from the full data base and to track different loan cohorts over time.

The data covers the time period from January 2002 until June 2010 at a monthly frequency and the analysis is focussed on loans originated between 2002 and 2008. I restrict the sample to prime, first, fixed-rate, 30-years mortgages that have a standard amortization schedule (are not balloon mortgages). I focus on only one mortgage type because the structural model would have to be recomputed for each different mortgage contract. The selection is motivated by the fact these are the most common mortgage contracts. The data base contains around 23 million loans with these characteristics in 2010.<sup>3</sup>

I further focus the analysis on loans with a loan-to-value ratio (LTV) above 95%, which depending on the year represents about 20 – 30% of all loans that satisfy the above restrictions. Looking at loans with different LTVs separately allows to generate a more accurate home equity distribution in the model. This is important due to the highly non-linear relationship between default decisions and negative equity. Furthermore, the loans with a high LTV default most frequently, so it makes sense to focus an analysis of mortgage default on them. But the main reason for concentrating on this group is a data problem. In the LPS data only the LTV of the first mortgage is observed, but

---

<sup>3</sup>Amromin and Paulson (2009) estimate that the LPS data covers about 60% of the prime market between 2004 and 2007. Elul, Souleles, Chomsisengphet, Glennon, and Hunt (2010) report that the LPS data covers about 70% of all mortgage originations in 2005 and 2006. But coverage varies by year with lower coverage in earlier years.

not the combined LTV of the first and a possible second mortgage.<sup>4</sup> Since the combined mortgage amount should be relevant for a borrower's decision to default the fact that second mortgages are unobserved is a problem for empirical work. This is a particular concern for structural models because of the strong role that theoretical approaches place on negative equity. In order to mitigate this data problem I thus focus on first mortgages with a very high LTV because these borrowers should be least likely to have a second mortgage on their home. However I also investigate whether and how the conclusions of the paper generalize to loans with a LTV of the first mortgage between 75% and 84% in section 9.

The data set contains aggregate information on contract characteristics by month of origination like the mortgage rate or credit scores. Furthermore, aggregate statistics on payment behavior are observed each month and broken down by the age of the loan. This allows to track the payment behavior of different cohorts of loans (defined by month of origination) over time. Specifically, each of these cells (defined by time period and loan age) contains how many active loans are delinquent or in foreclosure and how many are terminated through foreclosure or prepayment. Following much of the recent empirical literature cited in the introduction, I define a loan to be in default when it is 60 days or more past due, i.e. two payments have been missed. Accordingly, cumulative default rates for a loan cohort are constructed as the share of active loans that are 60 days or more delinquent or in foreclosure normalized by the share of initial loans that are still active plus the share of initial loans where foreclosure has already been completed.

## 2.2 House Prices

Information on house prices comes from the Federal Housing Finance Agency (FHFA). The monthly national and census division level repeat-purchase house price indices between 1991 and 2010 deflated by the Consumer Price Index (CPI) are used as measures of aggregate real house price movements. Estimates of the moments of the microeconomic house price distribution within a census division around the respective aggregate trend are used to generate a realistic house price distribution in the simulation. This is important because otherwise theoretical models cannot explain any default during times of positive aggregate house price growth.

---

<sup>4</sup>Elul, Souleles, Chomsisengphet, Glennon, and Hunt (2010) provide evidence that second mortgages are frequent and significantly affect the combined loan-to-value ratio. They report that on average 26% of all borrowers have a second mortgage and this adds on average 15% to the combined LTV. Unfortunately, they do not report a break-down of these statistics by the LTV of the first mortgage.

Throughout the paper the evolution of the real house price  $P_{it}$  of an individual house  $i$  in period  $t$  is modeled as

$$\ln(P_{it}) = \ln(P_{i,t-1}) + g_t^{agg} + g_{it}^{ind} \quad (1)$$

where the house price growth rate has two components, an aggregate component  $g_t^{agg}$  that is common to all houses and an individual component  $g_{it}^{ind}$  specific to the individual house. Such a formulation is consistent with the approach used by the FHFA to estimate the house price index, cf. the description in Calhoun (1996).<sup>5</sup> The general aim is to base the simulation framework as directly as possible on the information and empirical procedures of the FHFA.

In equation (1) a census division index was suppressed for convenience. But the aggregate trend represented by  $g_t^{agg}$  and the moments of  $g_{it}^{ind}$  are in fact specific to the census division in which the house is located. Thus, this paper uses data at the census division level and information on the regional composition of loan cohorts in the mortgage data. When drawing house prices the simulation draws are allocated across census divisions such that in each cohort the simulated sample has the same regional composition as in the mortgage data. The aggregate component  $g_t^{agg}$  represents the growth rate of the census division real house price index. In the simulation this component is taken directly from the data.

The individual component  $g_{it}^{ind}$  is unobserved. But the FHFA provides estimates of the variance that are used to simulate a realistic microeconomic house price distribution. Specifically, it is assumed that the individual component  $g_{it}^{ind}$  is independent over time and individuals and normally distributed with mean zero and variance  $V_t$ . The variance of  $g_{it}^{ind}$  depends on the time since the house was bought. This is a realistic feature of the data and based on estimates of the FHFA. Using my own notation, cf. footnote 5, the FHFA specifies a quadratic formula in time for the variance of the total individual part of the house price change since purchase given by

$$\text{Var} \left( \sum_{\tau=1}^t g_{i\tau}^{ind} \right) = \frac{\kappa}{3}t + \frac{\lambda}{9}t^2. \quad (2)$$

---

<sup>5</sup>I use a slightly different notation relative to the FHFA because I want to use this equation in a dynamic optimization problem and simulations. In order to see how it is related, rewrite equation (1) as

$$\ln(P_{it}) = \ln(P_{i,0}) + \sum_{\tau=1}^t g_{\tau}^{agg} + \sum_{\tau=1}^t g_{i\tau}^{ind}$$

where  $\ln(P_{i,0}) + \sum_{\tau=1}^t g_{\tau}^{agg} = \beta_t + N_i$  and  $\sum_{\tau=1}^t g_{i\tau}^{ind} = H_{it}$  give equation (1) in Calhoun (1996).

where an adjustment has been made for the fact that this paper operates at a monthly instead of a quarterly frequency. By the independence assumption the variance of  $g_{it}^{ind}$  is then given by

$$V_t = \text{Var} (g_{it}^{ind}) = \text{Var} \left( \sum_{\tau=1}^t g_{i\tau}^{ind} \right) - \text{Var} \left( \sum_{\tau=1}^{t-1} g_{i\tau}^{ind} \right) = \frac{\kappa}{3} + \frac{\lambda}{9}(2t - 1).$$

The FHFA provides estimates of  $\kappa$  and  $\lambda$  at the census division level that I use to generate realistic distributions around the division level aggregate trends. The estimates of  $\kappa$  are positive and those of  $\lambda$  are negative and small in absolute magnitude. This implies that the variance of  $\sum_{\tau=1}^t g_{i\tau}^{ind}$  increases less than linearly with time and the variance of a single  $g_{it}^{ind}$  is decreasing over time. On average across census divisions the estimates of  $\kappa$  and  $\lambda$  imply that the shock in the first month  $g_{i1}^{ind}$  has a standard deviation of about 2.49%, while after five years the standard deviation of  $g_{i60}^{ind}$  is around 2.37%. Hence the standard deviation of  $g_{it}^{ind}$  decreases relatively slowly over time.

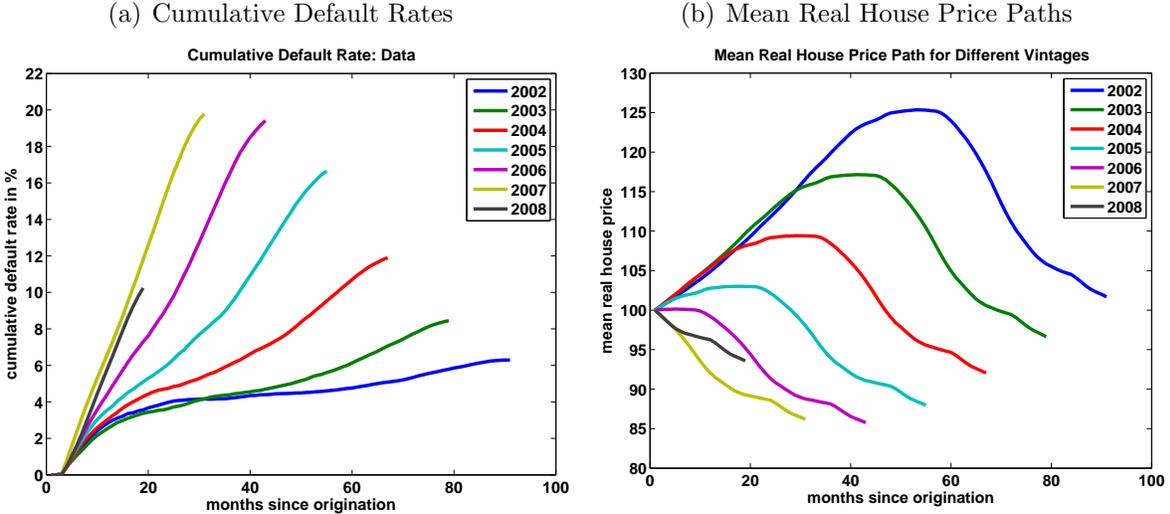
### 2.3 Empirical Facts on Default Rates and House Prices

The key empirical facts on mortgage default rates and house prices are presented in figure 1. Figure 1(a) shows the average cumulative default rates for loan cohorts originated between 2002 and 2008 grouped by the year of origination in my data set. The data shows clearly that loan cohorts originated later during this period defaulted much more frequently at the same time since origination. This increase constitutes part of the US mortgage default crisis and shows that the rise in default rates was in no way restricted to the subprime market and adjustable rate or hybrid mortgages.

Figure 1(b) presents the mean real house price paths (normalized to 100 at origination) for the cohorts of loans originated between 2002 and 2008. Borrowers of loans originated between 2002 and 2005 experienced on average rising real house prices during the immediate time after origination and falling house prices later during the course of the loan. In contrast, the real value of a home of a borrower who took out a mortgage between 2006 and 2008 decreased sharply immediately after origination.

The key research questions of the paper are motivated by the facts in figure 1. Can the variation in house price paths quantitatively explain the variation in mortgage default rates across cohorts within a structural economic model? What features should such an

Figure 1: Cumulative Default Rates and House Prices for different Origination Years



empirically successful model have? Does this variation allow to discriminate between different theoretical models of mortgage default?

### 3 Reduced Form Models

This section presents evidence on mortgage default from estimating and simulating two highly stylized models. But these models are motivated by economic theory and represent the simplest possible reduced forms of a frictionless option-theoretic model and the double-trigger hypothesis. The aim is to discriminate between these different theories in a relatively general way that is independent of the exact specification of the respective structural model. Building on these results the following section then develops a structural economic model that has hope to be empirically successful.

#### 3.1 Model Setup

The paper considers individual borrowers who took out a fixed-rate 30-years mortgage. Each loan cohort defined by origination date consists of many borrowers who are indexed by  $i = 1, \dots, N$  and observed in periods  $t = 1, \dots, T$  after loan origination. Borrowers take a single decision each period and can either service the mortgage or default on the loan and “walk away” from the house. Denote the default decision of an individual borrower  $i$  in month  $t$  after origination by a set of dummy variables  $d_{it}$ . The variables

$d_{it}$  take the value 1 once the borrower has defaulted, and the value 0 in all periods prior to default. Thus it is sufficient to present default decision rules in period  $t$  for situations when the borrower has not defaulted yet.

The next two sections present the two models, the “threshold” and “shock” model. Both models view negative equity as a necessary, but not sufficient, condition for default. The individual decision rules in the two models differ in how default exactly depends on house equity, and hence the house price and the mortgage balance.

For a fixed-rate mortgage the nominal mortgage balance  $M_{it}$  of borrower  $i$  evolves deterministically over time according to

$$M_{i,t+1} = (1 + r^m)M_{it} - m_i \quad (3)$$

where  $r^m$  is the monthly mortgage rate which is constant across individuals.  $m_i$  are fixed nominal monthly payments covering mortgage interest and principal. These payments are determined at the beginning of the contract and satisfy

$$m_i = \left[ \sum_{t=1}^T \frac{1}{(1 + r^m)^t} \right]^{-1} M_{i0} \quad (4)$$

where  $M_{i0}$  is the initial loan amount and the loan has a maturity of  $T = 360$  months. The initial loan amount is a function of the initial loan to value ratio  $LTV_i$  and initial house price  $P_{i0}$  and given by  $M_{i0} = LTV_i \times P_{i0}$ . Here borrowers are heterogenous with respect to the LTV. It is assumed that agents take decisions based on real variables. Thus it is useful to define the real mortgage balance as  $M_{it}^{real} = \frac{M_{it}}{\Pi_t}$  where  $\Pi_t$  is the CPI and  $\Pi_0 = 1$ . This assumption does not affect the results and the conclusions are identical when decisions are based on nominal variables.

The real house price  $P_{it}$  of an individual homeowner evolves according to equation (1). House price growth has an aggregate and individual component as described in section 2.2.  $P_{i0}$  is normalized to 100. This involves no loss of generality as seen below.

Due to the simplicity of the presented models I also add the constraint that default is only allowed from the fourth month since origination onwards. This is completely ad-hoc, but provides a better fit of both models to the data in the early periods after origination when default rates are essentially zero. But the comparison across models and the conclusions drawn below do not depend on this assumption.

### 3.2 The Threshold Model

The first model assumes that borrowers with negative equity default on their mortgage at the first time that the real value of equity falls below a certain threshold value. Therefore I call this the “threshold model”. Here, I adopt the simplest possible specification with a threshold that is proportional to the initial house price and constant over time given by  $\phi P_{i0}$  where  $\phi < 0$ . If in period  $t \geq 4$  the borrower has not defaulted yet then the default decision in that period is described by

$$d_{it} = \begin{cases} 1, & \text{if } P_{it} - M_{it}^{real} < \phi P_{i0} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

This is a simple reduced-form of a frictionless option model. The corresponding structural model would derive the threshold parameter  $\phi$  from optimizing behavior. For example the borrower might trade off the expected future capital gains on the house for the mortgage payments in excess of rents. Here I remain agnostic about the exact trade-off and the value of  $\phi$  and instead estimate it from the data.

### 3.3 The Shock Model

The second model assumes that borrowers with any level of negative equity only default on their mortgage when they also receive a default shock in that period. I call this the “shock model”. Again I adopt the simplest possible specification. The probability to receive a default shock  $\psi$  is constant and satisfies  $0 \leq \psi \leq 1$  and default shocks are independently and identically distributed over time. If the borrower has not defaulted yet, the default decision in period  $t \geq 4$  is determined by

$$d_{it} = \begin{cases} 1, & \text{if } P_{it} - M_{it}^{real} < 0 \text{ and the default shock occurs} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

This is a reduced-form of a double-trigger model. Here the default shock represents the life event like unemployment or divorce that combined with negative equity triggers default. The parameter  $\psi$  represents the probability that the life event occurs. Again  $\psi$  needs to be estimated from the data.

### 3.4 Model Simulation, Estimation and Test

Conditional on the respective model parameters  $\phi$  and  $\psi$  both models can be simulated for subsequent cohorts of loans originated each year between 2002 and 2008. For each cohort I draw 100,000 individual histories of house prices and default shocks with the same length in months as the respective cohort is observed in the data.<sup>6</sup> When computing the mortgage balance the mortgage rate is kept constant within a cohort and set equal to the respective cohort average. But borrowers within a cohort are heterogenous with respect to the LTV which varies in steps of one percentage point between 95% and 104%.<sup>7</sup> The frequency of these different loan-to-value ratios at origination is varied across cohorts as observed in the mortgage data. This means that possible changes to the average mortgage rate and the LTV distribution across cohorts are taken into account in the simulation. Data on the path of inflation rates from the CPI is used to compute the real mortgage balance. The decision rules are then applied to these shock histories and paths of the real mortgage balance.

The idea of the test procedure is to use only the default data of the cohort originated in 2002 to estimate the unknown parameters  $\phi$  and  $\psi$ . The test of the models is then based on out-of-sample predictions. Conditional on the parameter values estimated from the 2002 cohort, default rates for the cohorts 2003 to 2008 are simulated from the models. The test constitutes in comparing simulated and empirically observed default rates and checking which estimated model gives a better fit to the data.

The model parameters are estimated by a simulated method of moments procedure. Let  $\theta$  stand in for the parameter to be estimated in the respective model. The idea of the estimation is to choose  $\theta$  such that the cumulative default rates for the 2002 cohort simulated from the model match as well as possible those observed in the data. Collect the variables  $d_{it}$  in one vector  $D_i = [d_{i1}, \dots, d_{iT}]'$  for each individual. The mean of this vector  $\bar{D} = \frac{1}{N} \sum_{i=1}^N D_i$  represents the empirically observed cumulative default rate. The expected value of  $D_i$  is  $E[D_i] = D(\theta)$  and denote the expected value evaluated by simulation of  $S$  individuals from the model by  $\tilde{D}(\theta)$ . The deviation of the model from the data is then given by  $G(\theta) = \bar{D} - \tilde{D}(\theta)$ . The simulated method of moment estimator of  $\theta$  minimizes  $G(\theta)'WG(\theta)$  where  $W$  is a weighting matrix. I weight all moments equally

---

<sup>6</sup>The simulation procedure for individual house prices is explained in detail in section 2.2. For the shock model I also draw histories from an i.i.d. uniform distribution on the interval  $[0, 1]$ . For a given parameter  $\psi$  the default shock occurs for the respective individual and month if the uniform draw is smaller or equal to  $\psi$ .

<sup>7</sup>The few loans with a LTV above 104% are subsumed under the 104% LTV group.

by using an identity matrix as the weighting matrix.  $\theta$  is then estimated by minimizing a least squares criterion function given by

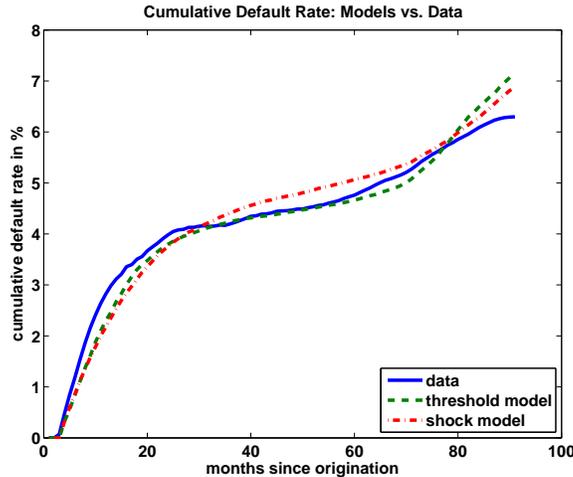
$$\frac{1}{T} \sum_{t=1}^T \left( \bar{d}_t - \tilde{d}_t(\theta) \right)^2 \quad (7)$$

where  $\bar{d}_t$  and  $\tilde{d}_t(\theta)$  are the  $t$ -th element in the vectors  $\bar{D}$  and  $\tilde{D}(\theta)$ , respectively. Here  $\tilde{d}_t(\theta)$  is evaluated using a frequency simulator such that  $\tilde{d}_t(\theta) = \frac{1}{S} \sum_{j=1}^S \tilde{d}_{jt}(\theta)$  and  $\tilde{d}_{jt}(\theta)$  represents the outcome for period  $t$  of applying the decision rules to the drawn history  $j$  of the underlying shocks. The minimization problem is solved by a grid search algorithm.

### 3.5 Results

For the threshold model the negative equity default threshold  $\phi$  is estimated as  $-11.0\%$ . This means borrowers default as soon as they have a real value of negative equity of  $11\%$  of the initial house price. In contrast, for the shock model the default shock probability  $\psi$  is estimated to be  $1.3\%$  such that each period  $1.3\%$  of those borrowers with negative equity default on their loan. The fit of the two models to the cumulative default rate of the 2002 cohort is shown in figure 2. Both models are able to fit this data very well.

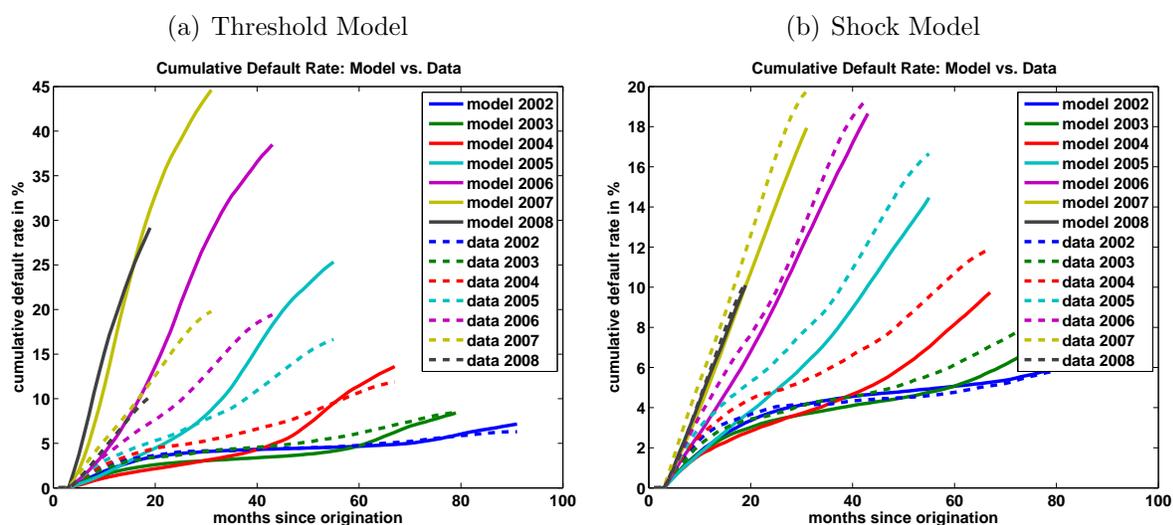
Figure 2: Cumulative Default Rate for 2002 Cohort: Models vs. Data



The next step is to test the two estimated models by checking how well they perform in predicting out-of-sample. Figure 3(a) shows the fit of the threshold model to the full sample of all cohorts between 2002 and 2008. The equivalent fit of the shock model is

presented in figure 3(b). It turns out that the threshold model has severe empirical problems. When it is forced to match default rates of the 2002 cohort, it over-predicts default rates for the later cohorts in the simulation period by at least one order of magnitude. The threshold model is excessively sensitive to the shifts in the mean of the house price distribution observed in the data. In contrast, the shock model gives a very good fit to the broad dynamics in the data. However, the shock model predicts too few defaults especially for the 2004 cohort and to some extent also for the 2003 and 2005 cohorts. This could imply that these cohorts were in fact composed of somewhat more risky borrowers though they appear to be similar based on observed characteristics discussed later in section 7.

Figure 3: Cumulative Default Rates for Loans originated in 2002-2008: Models vs. Data



The explanation for the difference between models is the following. The shock model predicts that a fraction  $\psi$  of borrowers with negative equity default each period. When the whole equity distribution shifts left due to the fall in aggregate house prices, the shock model predicts that the default rate should increase in proportion to the number of borrowers who experience negative equity. It turns out that observed default rates exhibit this pattern. But the threshold model is concerned with the (far left) tail of the equity distribution. It predicts that all borrowers with an extreme level of negative equity below  $\phi$  times the initial house price default. When the equity distribution shifts left the number of borrowers with such an extreme level of negative equity increases faster than the observed default rate. This generates the inconsistency with the data.

Two conclusions can be drawn from these results. First, an empirically successful structural model cannot rely on a single-trigger or threshold mechanism alone. Instead some

shocks other than house price shocks must play a role. Second, with a double-trigger model accounting for the increase in the fraction of borrowers with negative equity caused by the mean shift in house prices is sufficient to explain the broad rise in default rates. Together with the evidence on the stability of loan characteristics presented in section 7 this supports a hypothesis featuring a strong explanatory role of the macroeconomic house price movements and against the pool of borrowers becoming more risky per se.

Motivated by these results, the next main section presents a structural model featuring idiosyncratic unemployment risk and liquidity constraints. This serves several purposes. First, the model aims at providing micro-foundations for the double trigger hypothesis. This means to provide conditions under which a rational agent exhibits double-trigger default behavior. Second, it allows to check whether unemployment shocks can quantitatively play the role of the trigger events. One can also check whether the strong explanatory role of aggregate house prices survives in such a structural framework. Third, such a model can be used for policy analysis.

### **3.6 Robustness Checks**

This section reports a battery of robustness checks that were performed to scrutinize these results. I find that the results are robust across all the modifications considered here. Graphs equivalent to figure 3 for each of the performed scenarios are available upon request.

Instead of estimating the models on the 2002 cohort with low default rates, I also estimate them on the 2008 cohort with very high default rates. This does not affect the good fit of the shock model. But now the threshold model greatly undershoots the default rates of early cohorts and also still overshoots the 2006 and 2007 cohort. Thus the comparison across models is unaffected.

Another robustness check replaces the out-of-sample test with an in-sample test. Here I estimate the two models on all cohorts and then examine the fit within that sample. The threshold model still has considerable problems to match the data. It generally undershoots earlier cohorts and the early months after origination for all cohorts and at the same time still overshoots the late months of the 2006 and 2007 cohorts. In contrast, the shock model gives an excellent fit to the data. The conclusions across models are essentially unchanged.

I also examine the role of the variation in mortgage rates and the distribution of loan-to-value ratios across cohorts in three alternative specifications. In the first specification, I keep the within cohort LTV distribution fixed across cohorts according to the average frequency. The second specification abstracts from within cohort heterogeneity such that everyone has the same LTV according to the respective within cohort average. The third specification is the same as the second except that the LTV and mortgage rate are not varied across cohorts. All these changes have very modest effects on both models and leave the conclusions across models unaffected. This implies that the double-trigger model attributes the rise in default rates to the variation in aggregate house prices and not the changes in contract characteristics across cohorts. It also suggests that abstracting from this heterogeneity across cohorts in the structural model is not too restrictive.

In section 2.2 it was assumed that the individual house price shocks are normally distributed. The major argument supporting this choice is that by the central limit theorem the sum of individual shocks converges asymptotically to a normal distribution anyway. But since the analysis also covers periods where  $t$  is still small, I perform an additional check here. Instead of using a normal distribution for the individual shocks I specify them as being uniformly distributed on the interval  $[-b_t, b_t]$ . The parameter  $b_t$  is then chosen such that the variance of the uniformly distributed shock in period  $t$  in the respective census division is identical to the one used in the standard framework. I find that the results are almost identical.

Another potential concern is that the simplicity of the presented reduced-form models with only one constant parameter somehow biased the results against the frictionless option model. There is also no strong reason why the default threshold parameter  $\phi$  and default shock probability  $\psi$  should be constant over the course of a loan. It turns out that the results are robust to changing this assumption. As a check I have performed a scenario where the respective default parameter depends fully on the month since origination  $t$ . The constant parameters in the model are then replaced with  $\phi_t$  and  $\psi_t$  that are allowed to differ each period from  $t = 1, \dots, T$  when fitting the models to the 2002 cohort. Under these circumstances both models use all degrees of freedom of the data and perfectly match the 2002 cohort. The cumulative default rates simulated for the other cohorts then inherit the non-smoothness of the first differences of the cumulative default rate of the 2002 cohort. But subject to that qualification the conclusions on the out-of-sample fit remain essentially unchanged. The threshold model still greatly overshoots. The shock model generates default rates of the right magnitude, but predicts slightly lower default rates for some months compared to the benchmark specification.

## 4 Structural Model

This section introduces a theoretical model of the repayment decision of a homeowner who financed the home purchase with a fixed-rate mortgage. Each period the borrower chooses non-housing consumption and whether to stay in the house and service the mortgage or leave the house and terminate the mortgage. The mortgage can be terminated either by selling the house and repaying the mortgage or defaulting on the loan by "walking away". The homeowner faces uncertainty on the future price of the house, unemployment shocks and a borrowing constraint for unsecured credit. One period corresponds to one month. Throughout this section an individual index  $i$  is suppressed for convenience.

### 4.1 Mortgage Contract

The household took out a fixed rate mortgage with outstanding nominal balance  $M_0$  and nominal mortgage rate  $r^m$  to finance the purchase of a house of price  $P_0$  in period 0. Mortgage interest and principal have to be repaid over  $T$  periods in equal instalments of nominal value  $m$  that are fixed at the beginning of the contract and satisfy equation (4). Over time the outstanding nominal mortgage balance  $M_t$  evolves according to equation (3) as long as the household services the mortgage.

### 4.2 Preferences and Choices

Preferences are specified as in Campbell and Cocco (2003), but allow for a direct utility benefit of owning a house. Household decisions over the length of the mortgage contract are determined by maximizing expected life-time utility given by

$$U = \mathbb{E}_0 \sum_{t=1}^T \beta^{t-1} \left( \frac{C_t^{1-\gamma}}{1-\gamma} + \theta \mathcal{I}(\text{own}_t) \right) + \beta^T \frac{W_{T+1}^{1-\gamma}}{1-\gamma} \quad (8)$$

which is derived from consumption  $C_t$  in periods 1 to  $T$  and remaining wealth  $W_{T+1}$  at the end of the contract. The flow utility function is assumed to be of the CRRA form where  $\gamma$  denotes the parameter of relative risk aversion and the inverse of the intertemporal elasticity of substitution.  $\beta$  is the time discount factor.  $\mathcal{I}(\text{own}_t)$  is an indicator variable that is one if the agent owns a home in period  $t$  and zero otherwise.  $\theta$  is a direct utility benefit from being a homeowner reflecting for example an emotional attachment to the house. The specification of the utility function implicitly assumes that consumption and the size of the house are separable in the homeowner's utility function.

In each period the homeowner has to decide how much to consume and on staying or leaving the house. If the agent wants to leave this can be done by either selling the house (and repaying the current mortgage balance) or defaulting on the loan by "walking away".<sup>8</sup> It is assumed that a homeowner who leaves the house will rent a house of the same size for the rest of life.

### 4.3 Constraints

The dynamic budget constraint depends on the borrower's house tenure choice. For a homeowner who stays in the house it is given by

$$A_{t+1} = (1 + r)(A_t + Y_t - \frac{m}{\Pi_t} + \tau r^m \frac{M_t}{\Pi_t} - C_t) \quad (9)$$

where  $A_t$  denotes real asset holdings and  $Y_t$  real net labor income in period  $t$ . The real interest rate on savings  $r$  is assumed to be constant over time.  $m$  is the nominal payment to service the mortgage. But the nominal mortgage interest  $r^m M_t$  is tax deductible and  $\tau$  is the tax rate. All nominal variables need to be deflated by the current price level for consumption goods  $\Pi_t$  to arrive at a budget constraint in terms of real variables. The presence of  $\Pi_t$  generates the "mortgage tilt effect". This means that due to inflation the real burden of the mortgage is highest at the beginning of the contract and then declines over time. It is assumed that the inflation rate  $\pi$  is constant over time and  $\Pi_t$  thus evolves according to  $\Pi_{t+1} = (1 + \pi)\Pi_t$ .

In case the house is sold at the current real price  $P_t$ , the homeowner needs to repay the current outstanding nominal mortgage balance  $M_t$  and can pocket the rest. The budget constraint then reads as

$$A_{t+1} = (1 + r)(A_t + Y_t - R + P_t - \frac{M_t}{\Pi_t} - C_t). \quad (10)$$

Here  $R$  is the real rent for a property of the same size. It is assumed that an agent who terminates the mortgage through prepayment or default needs to rent for the rest of life. Real rents are assumed to be proportional to the initial house price and then constant over time as

$$R = \alpha P_0. \quad (11)$$

---

<sup>8</sup>The model does not include a mortgage termination through refinancing for computational reasons. Otherwise the mortgage balance becomes a separate state variable. This is unlikely to be a major limitation because refinancing is only feasible when the borrower has positive equity in the house. Thus it does not directly compete with the default decision in a negative equity situation.

This specification involves both a highly realistic feature of rents and an approximation. The realistic feature is that during the period of study real rents remained almost constant, while real house prices first increased and then decreased enormously. The specification implies that after origination the rent-price ratio decreases when real house prices increase. Such a negative relationship between the rent-price ratio and real house prices exists in the data provided by Davis, Lehnert, and Martin (2008) not only during the recent period, but at least since 1975. In this paper I take these observations as given and specify the exogenous variables of the model accordingly. But explaining this pattern is an important area for future research. However a fully realistic specification would also require to make  $\alpha$  cohort-specific. But I use an approximation for computational reasons such that  $\alpha$  is constant across cohorts and calibrated to a suitable average.

In contrast, if the agent decides to default on the mortgage by "walking away" or is already a renter the budget constraint is given by

$$A_{t+1} = (1 + r)(A_t + Y_t - R - C_t). \quad (12)$$

It is assumed that for reasons not explicitly modeled here the household faces a borrowing constraint for unsecured credit given by

$$A_{t+1} \geq 0. \quad (13)$$

Together with the budget constraints above this implies that the amount of resources available for consumption in a period depend on the house tenure choice.

Remaining wealth at the end of the contract for a homeowner is given by  $W_{T+1} = A_{T+1} + Y_{T+1} + P_{T+1}$  and for a renter by  $W_{T+1} = A_{T+1} + Y_{T+1}$ .

#### 4.4 Labor Income Process

The household's real net labor income  $Y_t$  is subject to idiosyncratic unemployment shocks and exogenously given by

$$Y_t = \begin{cases} (1 - \tau)Y_0 & \text{if employed} \\ \rho(1 - \tau)Y_0 & \text{if unemployed} \end{cases} \quad (14)$$

where  $Y_0$  is initial real gross income,  $\tau$  is the tax rate and  $\rho$  is the net replacement rate of unemployment insurance. Over time employment status evolves according to a Markov

transition process with the two states “employed” and “unemployed” and constant job separation and finding probabilities. Employed agents lose their job with probability  $s$  and stay employed with probability  $(1-s)$ . Unemployed agents find a job with probability  $f$  and stay unemployed with probability  $(1-f)$ .

There are several reasons why I focus on income fluctuations due to unemployment risk here. First, unemployment involves a severe fall in labor income from one month to another. This makes it a very plausible cause for short run liquidity problems. Second, other frequently used specifications of income processes as for example in Campbell and Cocco (2003) are typically calibrated for yearly frequencies. Thus, they are not directly applicable to a monthly framework. In any case, most of the income variation from month to month probably comes from unemployment spells and it therefore seems preferable to use such a process explicitly. Third, this allows to relate the model more closely to the double-trigger hypothesis and the empirical literature that has provided evidence that default is correlated with state unemployment rates.

I also abstract from deterministic changes to labor income like a life-cycle profile and keep the labor income of employed and unemployed agents constant over time. The reason is that I do not have any demographic information on the borrowers in my data set.

## 4.5 House Price Process

Real house prices are exogenous and evolve over time as specified in section 2.2 and equation (1). It is assumed that homeowners view the aggregate component  $g_t^{agg}$  of house price appreciation to be stochastic and distributed according to an i.i.d. normal distribution with mean  $\mu$  and variance  $\sigma^2$ . This process for the aggregate house price component is only used for forming agents’ expectations. In the simulation the realizations of  $g_t^{agg}$  are those observed in the data. For the individual component agents know that  $g_t^{ind}$  is distributed normally with mean zero and time-varying variances that depend on the parameters  $\kappa$  and  $\lambda$  as specified in section 2.2. In order to reduce the computational burden when computing policy functions the parameters  $\mu$ ,  $\sigma$ ,  $\kappa$  and  $\lambda$  are not varied across the nine census divisions. Instead they are set equal to national averages, cf. section 5.2 on the calibration. But the realizations in the simulation of the model of course come from the division specific data and distributions.

## 4.6 Initial Conditions

The homeowner solves the dynamic stochastic optimization problem conditional on initial asset holdings  $A_0$ , initial employment status, an initial loan-to-value ratio  $LTV = \frac{M_0}{P_0}$  and a debt to (gross) income ratio  $DTI = \frac{m}{Y_0}$ .<sup>9</sup> Without loss of generality, the initial house price  $P_0$  is normalized to 100.  $LTV$  and  $DTI$  then uniquely determine  $M_0$  and  $Y_0$ .

## 4.7 Computation

The borrower's optimization problem is characterized by four state variables (liquid wealth  $X_t = A_t + Y_t$ , employment status  $L_t$ , house price  $P_t$  and time  $t$ ) and two choice variables (consumption  $C_t$  and the mortgage termination choice). Note that for a fixed-rate mortgage the mortgage balance  $M_t$  evolves deterministically over time and is thus captured by the state variable  $t$ . The solution proceeds backwards in time. The continuous state and control variables are discretized and the utility maximization problem in each period is solved by grid search. Expected values of future variables are computed by Gaussian Quadrature. Between grid points the value function is evaluated using cubic interpolation.

## 4.8 Model Simulation

The model presented above is a dynamic stochastic partial-equilibrium model that maps contract characteristics at origination and realizations of the stochastic processes for house prices and employment status into default decisions. I simulate the model for subsequent cohorts of loans originated each month between January 2002 and December 2008 from the respective origination month until June 2010. For each cohort I draw 20,000 individual house price and employment histories with the same length in months as the respective cohort is observed in the data. House price histories are drawn as explained in section 2.2 and employment histories are drawn from the two-state Markov process specified in section 4.4.

Accordingly, within a cohort borrowers face the same aggregate house price movements (except for the differences between census divisions), but different individual house price and employment shocks. Differences between cohorts are generated from different paths of aggregate house prices depending on the date of origination.

---

<sup>9</sup>The name debt to income ratio is part of standard mortgage terminology, but can be easily misunderstood. It means the ratio of the monthly mortgage payment to gross income.

## 5 Parametrization

The structural model is parameterized in two steps. First the mortgage contract, house price expectations, rents, labor income, interest and inflation rates are calibrated to data on the respective variables, i.e. to data other than default rates. Then due to identification concerns the preference parameters are divided into a set that is calibrated ad-hoc and another that is estimated such that the model fits the cumulative default rates of the 2002 loan cohort. All parameter values are summarized in table 1 below. The model is solved at a monthly frequency. But a few parameters are presented at their yearly values if it is more convenient for comparison.

### 5.1 Contract Characteristics

This paper restricts attention to 30-years ( $T = 360$  months) fixed-rate mortgages. I use average characteristics at origination of the loans in my data set to determine the loan-to-value ratio, mortgage rate and debt-to-income ratio. The average initial loan-to-value ratio of these loans is 98.2%, so I set  $LTV = 98.2\%$ . The nominal mortgage rate  $r^m$  is set to 6.4% per annum which is the average mortgage rate for newly originated loans in my data set. The debt-to-income ratio  $DTI$  is set to 40% as in the data.<sup>10</sup> Naturally, all of these parameters could be changed in order to model different mortgage contracts.

### 5.2 House Price Expectations

As explained before, when computing policy functions the parameters  $\mu$ ,  $\sigma$ ,  $\kappa$  and  $\lambda$  are not varied across the nine census divisions. Instead they are set according to national averages in order to reduce the computational burden. The monthly house price index from the FHFA at the national level between 1991 and 2010 deflated by the Consumer Price Index (CPI) is used to estimate the parameters  $\mu$  and  $\sigma$  of the aggregate component. I find that at a monthly frequency  $\mu = 0.065\%$  and  $\sigma = 0.55\%$ . These values imply expected yearly aggregate real house price growth of 0.8% and a yearly standard deviation

---

<sup>10</sup>The data on the DTI is the only mortgage variable in the whole paper that is based on a somewhat different loan selection. The reason is that the DTI was not available in the tool that was used to aggregate and extract information from the LPS loan-level data set. Instead LPS provided me with a separate tabulation where it was not possible to use the same selection criteria. Specifically, the DTI information is for the same LTV class as the rest of the data, but it does not only cover prime, fixed-rate, 30-years mortgages. However the vast majority of loans in the LPS data are prime, fixed-rate mortgages and the modal maturity of these loans is 30 years, so this information should at least be a good approximation to the actual loan pool I consider.

of 1.9%. This calibration procedure implies that agents in the model have expectations on real aggregate house price growth that on average were correct in the years 1991 to 2010 as far as the mean and standard deviation are concerned.

The parameters  $\kappa$  and  $\lambda$  are determined as a simple average of the ones estimated by the FHFA for each of the nine census divisions. This gives  $\kappa = 0.00187$  and  $\lambda = -4.51E - 6$  and implies that the individual house price growth shock  $g_{it}^{ind}$  in the first month after house purchase is expected to have a standard deviation around 2.5%.

### 5.3 Income Process

The average tax rate  $\tau$  is set to 16% and the net replacement rate of unemployment insurance  $\rho$  to 62%. This is based on the OECD Tax-Benefit calculator for the United States. Specifically, the average loan amount, mortgage rate and debt-to-income ratio are used to determine the average gross income of the borrowers in the data set. Based on gross income the calculator reports the net income in work and out of work which then determine the average tax and net replacement rates. These calculations take taxes, social security contributions, in-work and unemployment benefits into account. Precise numbers especially for the tax rate also depend on the demographics of the household. I have used the average values for a married couple with one earner and no children.

Data from the Bureau of Labor Statistics on the national unemployment rate and median unemployment duration are used to compute time-series of monthly job finding and separation probabilities. This is done using steady state relationships. Since the data on median unemployment duration is reported in weeks, I first transform it to months by multiplying the weekly value by 12/52. The resulting median duration  $d$  in months is of course in general not an integer value. Given that I operate in discrete time I use an approximation to the relationship between median duration and the monthly finding probability  $f$  in steady state given by

$$(1 - f)^{\underline{d}-1} f + (d - \underline{d})(1 - f)^{\underline{d}} f = 0.5 \tag{15}$$

where  $\underline{d}$  is the next integer number lower than or equal to  $d$ . If the median duration in months is an integer value then the second term in equation (15) is zero. If it is not an integer value then the second term gives an approximation to the number of unemployed who find a job between month  $\underline{d}$  and  $\underline{d} + 1$  for a given finding probability  $f$ .

The steady state relationship between the unemployment rate  $u$  and job finding probability  $f$  and job separation probability  $s$  in the flows approach to unemployment is well known and given by

$$u = \frac{s}{s + f}. \quad (16)$$

Equations (15) and (16) are then solved for the time-series of  $f_t$  and  $s_t$  implied by the time-series of the unemployment rate  $u_t$  and median duration  $d_t$ .<sup>11</sup>

I then set  $s = 1.8\%$  and  $f = 31\%$  which are the average values of the computed monthly finding and separation probabilities for the period from 1990 to 2010. These values imply a steady state unemployment rate around 5.7%.

## 5.4 Other Prices

Nominal interest rates for 1-year Treasuries and changes to the Consumer Price Index (CPI) are used to compute real interest rates and inflation rates. Based on this data between 1990 and 2010 the real interest rate  $r$  is set equal to 1.4% and the inflation rate  $\pi$  equal to 2.7% on an annual basis. The initial rent-price ratio parameter  $\alpha$  is set equal to 4.0% on a yearly basis which is the average rent-price ratio between 2002 and 2008 in the data provided by Davis, Lehnert, and Martin (2008).

## 5.5 Initial Conditions

Initial assets and employment status are unobserved. But it seems reasonable that borrowers were employed when they got their loan, so I assume that. With respect to initial assets  $A_0$ , I use the computed policy functions to set initial assets equal to the buffer-stock desired by a borrower in period 1 who is employed and faces a house value equal to  $P_0$ . Thus I shut down possible effects from borrowers first converging to their desired buffer-stock and being more vulnerable to income shocks during the time immediately after origination.

---

<sup>11</sup>As a check on this procedure I predict the unemployment rate from the dynamic equation of unemployment  $u_{t+1} = u_t + s_t(1 - u_t) - f_t u_t$  using the computed time series of finding and separation probabilities as inputs. It turns out that this gives an excellent fit to the path of the actual unemployment rate.

Table 1: Model Parameters

Contract characteristics	Contract length in months	$T$	360
	Mortgage rate (yearly)	$r^m$	6.4%
	Initial loan-to-value ratio	$LTV$	98.2%
	Initial debt-to-income ratio	$DTI$	40%
House price process	Mean of aggregate component	$\mu$	0.065%
	Standard deviation of aggregate component	$\sigma$	0.55%
	Linear coefficient in individual component	$\kappa$	0.00187
	Quadratic coefficient in individual component	$\lambda$	-4.51E-6
Income process	Job separation probability	$s$	1.8%
	Job finding probability	$f$	31%
	Tax rate	$\tau$	16%
	Net replacement rate of unemployment insurance	$\rho$	62%
Other prices	Real interest rate (yearly)	$r$	1.4%
	Inflation rate (yearly)	$\pi$	2.7%
	Rent-price ratio (yearly)	$\alpha$	4.0%
Preferences	CRRA coefficient	$\gamma$	4
	Discount factor (yearly)	$\beta$	0.9
	Utility benefit of owning	$\theta$	0.18

## 5.6 Preferences

Ideally the three preference parameters  $\beta$ ,  $\gamma$  and  $\theta$  would all be estimated such that the model gives the best fit to the data on default rates. But it is well known that dynamic discrete choice models are not fully identified, cf. the discussion and references in Magnac and Thesmar (2002). Furthermore, given the complexity of the model estimating several parameters would be computationally costly. Faced with this situation I decide to calibrate the parameters  $\beta$  and  $\gamma$  ad-hoc and estimate only  $\theta$ . I also investigate how much the results depend on the specific choice of  $\beta$  and  $\gamma$ .

The parameters  $\beta$  and  $\gamma$  appear in most dynamic economic models and estimating them is the aim of a vast empirical literature. But unfortunately these empirical studies have not produced reliable estimates. For the discount factor  $\beta$  on a yearly basis the survey of Frederick, Loewenstein, and O'Donoghue (2002) shows that empirical estimates cluster over the full range between 0 and 1. For the intertemporal elasticity of substitution, which is the inverse of  $\gamma$ , Guvenen (2006) reviews empirical estimates ranging from around 1 to 0.1, which implies values of  $\gamma$  ranging from 1 to 10.

My impression is that many economists regard values of  $\beta$  below 1, but not too much

below 1, and values of  $\gamma$  between 1 and 4, possibly even up to 10, as reasonable. But strong views on specific parameter values are probably not warranted given the empirical evidence. The large variation in estimates could also reflect that preferences are not stable across choice situations and individuals. With respect to the intertemporal elasticity of substitution Guvenen (2006) argues that conflicting estimates can be reconciled if the rich have a high and the poor have a low elasticity. I follow his argument and since the average borrowers in my data set belong to the lower half of the income distribution, I set a relatively high value of  $\gamma = 4$ . This implies an intertemporal elasticity of substitution of 0.25. For  $\beta$  I choose a value of 0.9 at a yearly frequency in order to be below, but still close to 1. Compared with assumptions in many macroeconomic studies this might appear as a low value. But adapting Guvenen's argument to  $\beta$ , this does not necessarily conflict with other studies. The reason is that I am analyzing a particular pool of borrowers who are not rich and were only able to make a very small down-payment. This could be due to the fact that they are very impatient. The other agents in the economy who are net savers and lenders could then have a higher discount factor more in line with the macro literature. In any case these are only the benchmark values and I also investigate the sensitivity of the main results to these parameter choices.

Given values of  $\beta$  and  $\gamma$ , the preference parameter  $\theta$  representing the direct utility benefit from owning the house is estimated by the simulated method of moments. The procedure is identical to the one used earlier for the reduced-form models, cf. section 3.4. Again the parameter is chosen such that cumulative default rates simulated from the model match those observed in the data using only information from loans originated in 2002. This yields an estimated value for  $\theta$  of 0.18. The remaining data is used to test the ability of the estimated model to predict out of sample.

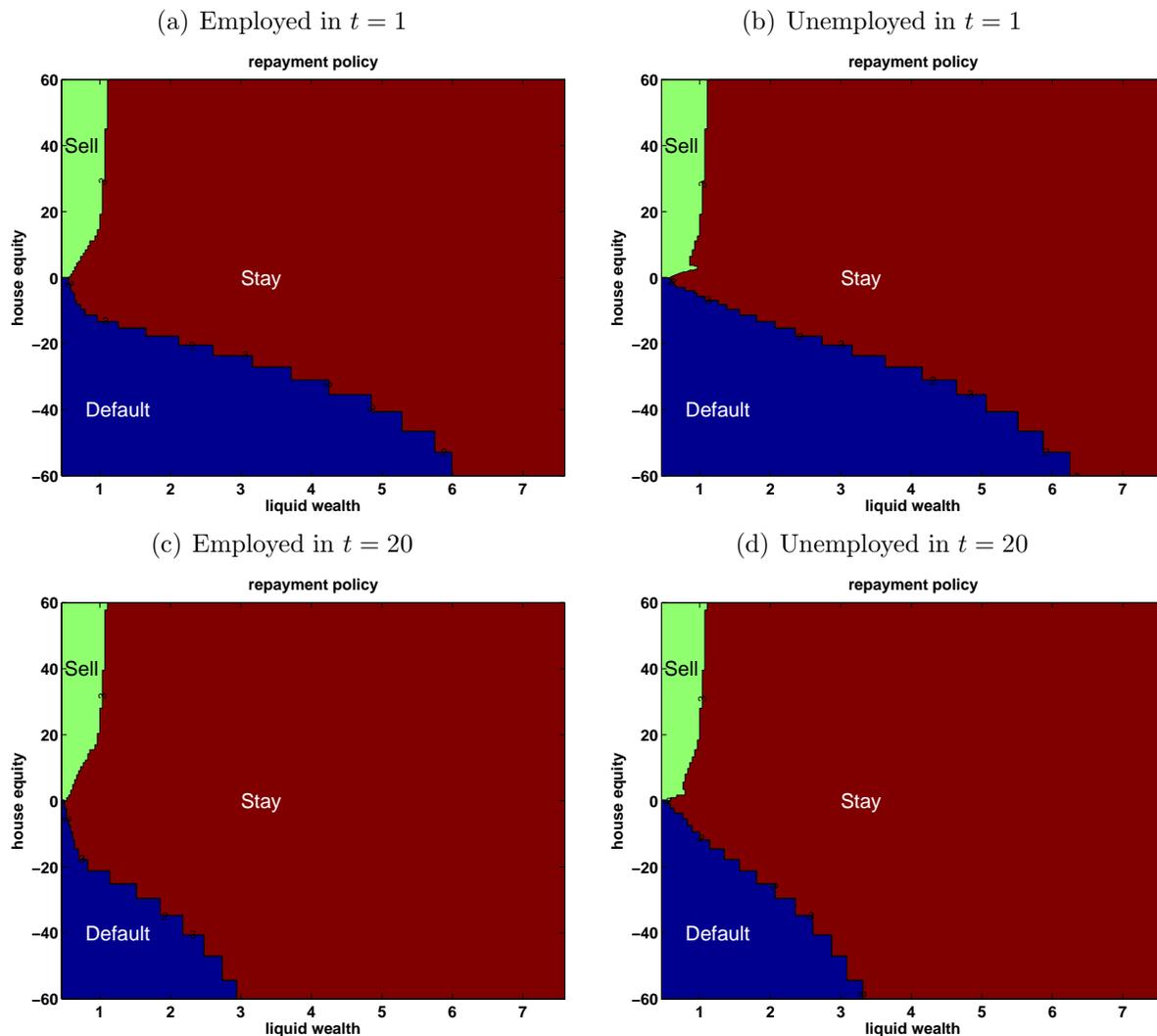
## 6 Results

This section explains the repayment policy function of a homeowner and the basic mechanism generating default over the life-cycle of a loan in the model. Then the main results how well the model fits the rise in default rates across loan cohorts are presented. Finally, a sensitivity analysis explores how the model depends on certain preference parameters.

## 6.1 The Repayment Policy Function

The repayment policy function of a borrower in the model is presented in figure 4 as a function of house equity, liquid wealth, employment status and time. Several features are note-worthy. First, negative equity is a necessary condition for default. Instead, with positive equity selling is strictly preferred to defaulting because the borrower is the residual claimant of the house value after the mortgage balance has been repaid.

Figure 4: Repayment Policy Function



*Notes:* Repayment choice as a function of the state variables liquid wealth, house equity, employment status and time. Blue region: Default. Green region: Sell. Red region: Stay.

Second, negative equity is not sufficient for default. There are many combinations of state variables where a borrower with negative equity prefers to stay in the house and service the mortgage. In a negative equity situation the basic trade-off of the borrower is the

following (postponing the role of the borrowing constraint until the next paragraph). The cost of staying in the house is that the borrower needs to make the mortgage payment, which is higher than the rent for an equivalent property. The benefit of staying is that the borrower receives the utility benefit of owning a house and keeps the option to default, sell or stay later. Specifically, there are possible future states of the world with positive equity. But the probability of reaching these states depends on the current house price. This establishes a default threshold level of the house price. Of course, when making this decision the rational borrower will also need to discount these future gains and take risk aversion into account.

Third and importantly, the level of negative equity at which the borrower exercises the default option depends on non-housing state variables: liquid wealth and employment status. Specifically, a borrower who is unemployed and/or has low liquid wealth will default at lower levels of negative equity. There are two reasons for terminating the mortgage in these states. One is that current borrowing constraints may bind and the borrower terminates the mortgage to increase current consumption. The other reason is that in these states it becomes very likely that borrowing constraints bind in the future and the agent is forced to terminate the mortgage then. But an anticipated future mortgage default creates an incentive to default already today to save the difference between the mortgage payment and the rent in the meantime. This also explains why unemployment, which is persistent, shifts the default frontier to the right.

Fourth, over time the default region shrinks. This is mainly due to the effect of inflation that diminishes the real difference between the effective mortgage payments and rents. This has two implications. First, a liquidity constrained borrower cannot increase current consumption much by a mortgage default. Second, staying in the home eventually dominates renting in all states because it simply becomes the cheaper option to live.

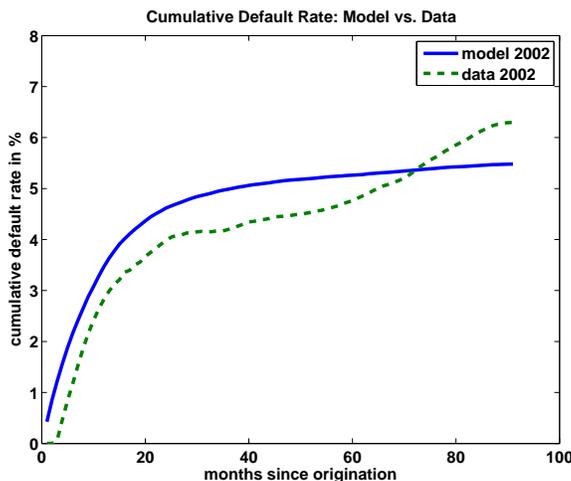
## **6.2 Default over the Loan Life-Cycle**

In this section I compare model results and data on the cohort of loans for which I have the longest time dimension in order to get an impression of default behavior over the life-cycle of a loan. Figure 5 presents the average cumulative default rate for loans originated in 2002. This is the cohort on which the model is estimated. Accordingly, the dynamics of default over the life-cycle of this cohort are captured relatively well by the model. But the model predicts too many defaults in the first months after origination and too few

in the very late months. I will discuss the reasons for this in more detail in the next section.

Though this cohort faces growing average house prices during the immediate time after origination as seen in figure 1(b), some individuals experience falling house prices and negative equity as a consequence of individual house price shocks. Households with negative equity default when prolonged stretches of unemployment have exhausted their buffer stock savings, cf. the default region of the state space in figure 4. Eventually, the cumulative default rate levels off due to two reasons. First, borrowers who are still active have amortized their mortgages sufficiently such that most have positive equity. Second, due to the mortgage tilt effect the difference between the real mortgage payment and real rents shrinks over time such that a default becomes less appealing.

Figure 5: Cumulative Default Rates of 2002 Cohort: Structural Model vs. Data

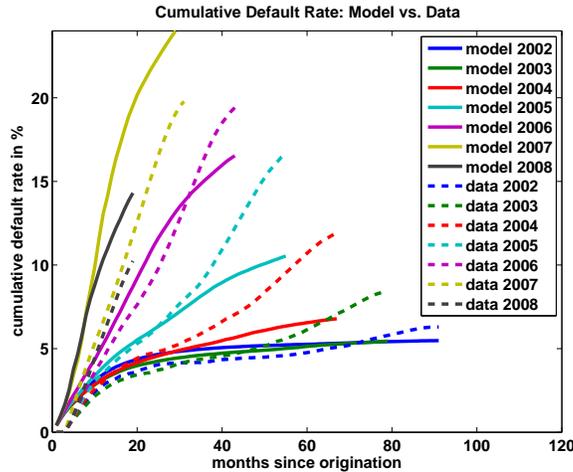


### 6.3 The Rise in Cumulative Default Rates

The next step is to compare the default behavior of different cohorts during the time period of the U.S. mortgage crisis. Figure 6 presents average cumulative default rates for cohorts of loans originated each year between 2002 and 2008.

When average house price appreciation slows down and eventually becomes negative as witnessed in figure 1(b) a higher fraction of borrowers experience negative equity which translates into more frequent default. The model can explain the broad pattern in the data and attributes the rise in cumulative default rates across cohorts to the different

Figure 6: Cumulative Default Rates of 2002-2008 Cohorts: Model vs. Data



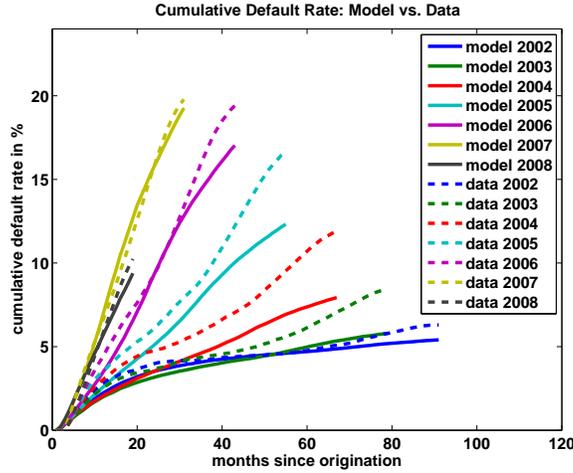
aggregate house price paths. The model is particularly successful in the early months after loan origination, but has problems to explain default in later months. In the model this is due to the effect of inflation, the mortgage tilt effect. This effect diminishes the difference between real mortgage payments and rents over time. The model is sensitive to this difference and reacts too strongly compared to the data. It is also noteworthy that the model inflation rate is constant and calibrated to the average inflation rate between 1990 and 2010 which is 2.7%. But in the final years of the simulation period inflation was much lower. For example on average between 2008 and 2010 it was 1.4% with 0.1% in 2008, 2.7% in 2009 and 1.5% in 2010. It is likely that the model would perform better for these actual inflation rates.

## 6.4 Role of Inflation

In this section I confirm that the role of inflation in the model and how I calibrated it are responsible for the poor performance of the model during periods long after origination. I simply change the inflation rate  $\pi$  ad-hoc to 1% instead of 2.7% in the benchmark calibration. All other parameters are unchanged, but  $\theta$  is reestimated at a value of 0.33 to fit the 2002 cohort. Figure 7 presents these results. The fit of the model improves and is now comparable to the one of the reduced-form double-trigger model, cf. figure 3(b).

There are at least two possible ways how to interpret the results in sections 6.3 and 6.4 on the role of inflation for the fit of the model. One possibility is that in the real world

Figure 7: Performance of the model for a low inflation rate



borrowers do not fully understand or underestimate the effect of inflation. This could be the reason why the model with a rational agent does not explain default so well in periods long after origination. It could also be that moving away from policy functions that are conditional on a constant inflation rate would improve the fit of the model.

The other possible interpretation is that unemployment and liquidity problems are not able to explain default in periods long after loan origination. Instead other reasons like marital break-up that were excluded from the structural model could be responsible for default in these periods. This paper only analyzes whether and how unemployment shocks could act as the trigger event in a structural model and found that they could definitely play an important role. But assessing the role of other life events and a decomposition of actual default rates into the different causes within the double-trigger paradigm is an important area for future research.

## 6.5 Dependence on Preference Parameters

All results from theoretical models depend in some way on parameters and the model presented here is no exception. Unfortunately, it is not easy to provide an exact characterization of the parameter space for which the agents in the model exhibit double-trigger default behavior because of the lack of a closed-form solution. But this section computes results for some examples of alternative parameter values for  $\beta$  and  $\gamma$  in order to get an idea how the model behaves in different parts of the parameter space.

The benchmark preference parameter values are  $\beta = 0.9$  and  $\gamma = 4$ . Here I consider all combinations of  $\beta \in \{0.85, 0.9, 0.95\}$  and  $\gamma \in \{2, 4, 6\}$ . For each of these  $(\beta, \gamma)$ -combinations the parameter  $\theta$  is reestimated in order to fit the 2002 cohort. All other parameters are as in the benchmark calibration. The resulting values of  $\theta$  for all combinations of  $\beta$  and  $\gamma$  are presented in table 2.

Table 2: Dependence of the estimated value of  $\theta$  on  $\beta$  and  $\gamma$

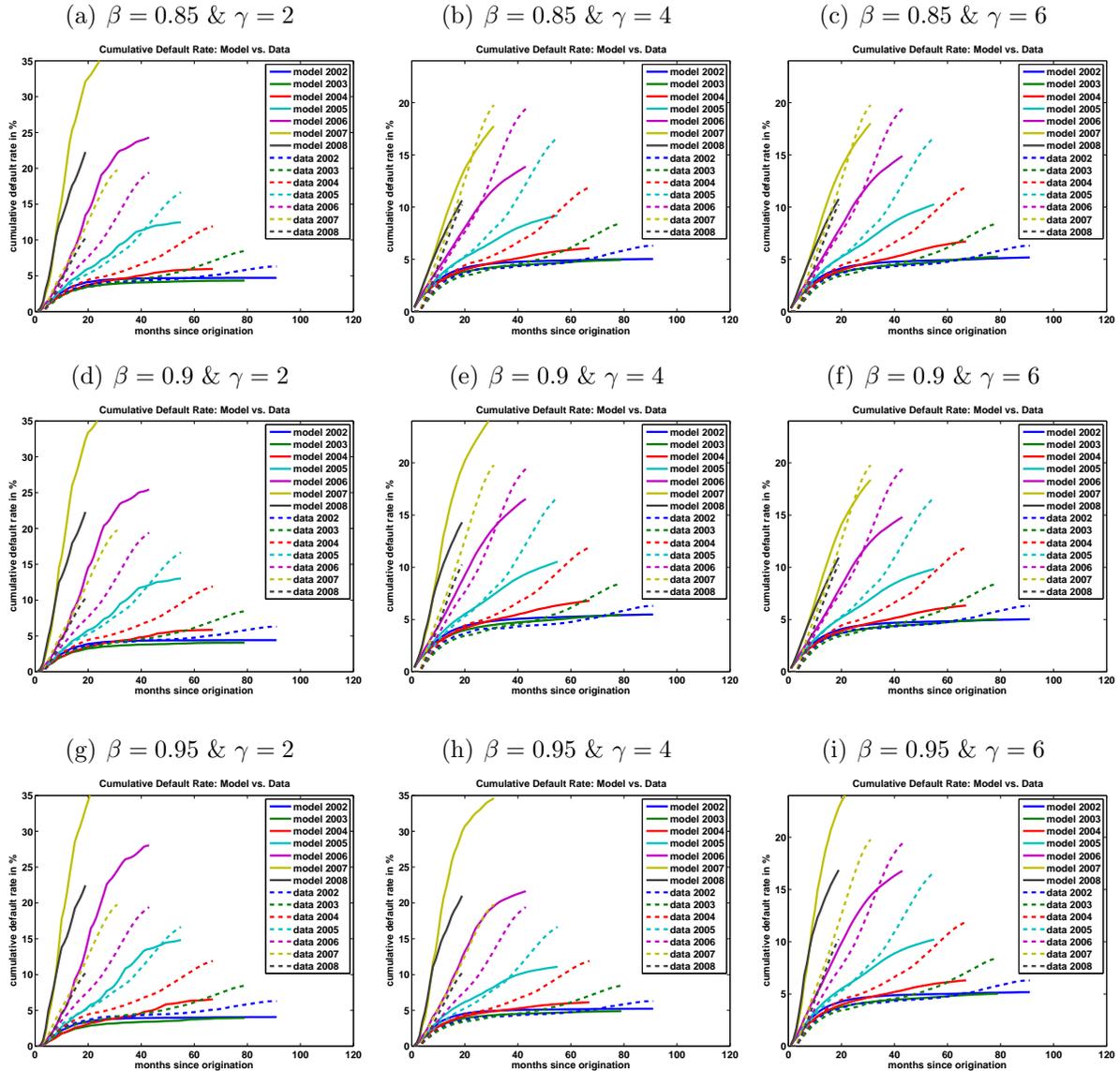
	$\gamma = 2$	$\gamma = 4$	$\gamma = 6$
$\beta = 0.85$	0.09	0.27	0.50
$\beta = 0.90$	0.05	0.18	0.36
$\beta = 0.95$	-0.02	0.07	0.20

The results for the different parameter combinations are presented in figure 8. The graphs are ordered such that  $\gamma$  increases horizontally from 2 (left) to 6 (right) and  $\beta$  increases vertically from 0.85 (top) to 0.95 (bottom). These results show that the model works as well or better than in the benchmark calibration for higher values of  $\gamma$  and/or lower values of  $\beta$ . These parameter changes make the agent less willing to substitute intertemporally and/or more impatient to consume today. This worsens the liquidity problem caused by unemployment. The model can only feature double-trigger behavior when being employed and being unemployed are sufficiently different. In contrast, for lower values of  $\gamma$  and higher values of  $\beta$  temporary income reductions can more easily be smoothed out. The model then implies that a sizeable portion of employed agents default in all cohorts. This brings the model close to a frictionless option model and the model then inherits all the problems of such a specification witnessed in section 3.

## 7 Discussion of an Alternative Explanation

All mortgage default theories hypothesize that default by a borrower is a function of the house price. This paper has presented further evidence that supports this view. However there is a competing explanation in the public and academic debate for the rise in default rates observed in figure 1(a). This explanation is that lending standards deteriorated sharply before the mortgage crisis. If this were true then the increase in mortgage default rates across cohorts could be due to a worsening of the loan quality. This would then also confound the empirical relationship between default rates and house prices that I use to

Figure 8: Sensitivity to Preference Parameters  $\beta$  and  $\gamma$



test mortgage default theories. Thus, this section presents evidence that loan quality is fairly stable across cohorts in my data set.

First of all I only look at data on prime fixed-rate mortgages. Therefore a shift towards more risky lending as far as it manifests itself in a shift from prime to subprime lending or from fixed to variable rate or hybrid mortgages is ruled out by construction. These compositional effects might or might not be significant contributors to the overall mortgage crisis, but they do not affect my analysis. We see clearly from figure 1(a) that even without such compositional effects mortgage default rates have increased substantially.

Another concern is that the loan-to-value ratio (LTV) might have increased over time leaving a smaller buffer before borrowers experience negative equity. I only consider loans that have a LTV above 95% and thus limit this possibility to shifts within that class of loans. Within this class the average LTV is basically constant across cohorts and only fluctuates mildly around the average value of 98.2% as seen in the first row of table 3. In the reduced-form models I even controlled for changes to the distribution of LTVs and found that the observed changes are irrelevant for the models considered here.

The second row of table 3 reports the average FICO credit score at origination of the different loan cohorts. These are very stable as well. To the extent that these credit scores are good measures of credit worthiness a significant deterioration in loan quality is not observable here.

Table 3 also contains information on the average mortgage rate that different cohorts face. A higher mortgage rate might make the loan as such less attractive to the borrower. There is some variation in this variable across cohorts. But the mortgage rate and default rates seem to be fairly uncorrelated across cohorts.

The average debt-to-income (DTI) ratio representing the share of the required mortgage payment in gross income is presented in the last row of table 3.<sup>12</sup> This has increased over time indicating that borrowers in later cohorts need to devote more of their gross income to service the mortgage. But the increase was by all means modest.

Table 3: Average Loan Characteristics at Origination by Loan Cohort

Cohort	2002	2003	2004	2005	2006	2007	2008	Average
Loan-to-value ratio in %	98.2	98.3	98.2	98.3	98.4	98.1	97.8	98.2
FICO credit score	676	673	669	670	668	670	678	672
Mortgage rate in %	6.9	6.0	6.1	6.0	6.6	6.7	6.2	6.4
Debt-to-income ratio in %	39	40	40	40	40	42	42	40

These statistics show that there is no strong evidence in favor of a deterioration of lending standards over time in my data set of prime fixed-rate mortgages with a LTV above 95%.<sup>13</sup>

<sup>12</sup>Footnote 10 also applies here.

<sup>13</sup>This conclusion might be specific to the prime market. For example Demyanyk and Van Hemert (2011) present evidence that loan quality deteriorated in the subprime market. But Amromin and Paulson (2009) also note that it is less obvious that a similar deterioration was present in the prime market. A particular advantage of my descriptive statistics is that they are based on all loans in the LPS data base satisfying my sample selection criteria. In contrast, other empirical studies using LPS data typically work with a 1% random sample such that their descriptive statistics are based on far fewer observations.

I conclude that this loan pool and time period indeed constitute a good testing ground for mortgage default theories.

One limitation of the paper is that it does not fully control for variation in contract characteristics across and within cohorts for computational reasons and the fact that I only have aggregate data. The evidence presented here suggests that this is not a major limitation because the different origination characteristics are quite stable. The reduced-form models also took variation of the mortgage rate and LTV distribution across cohorts into account and found that it cannot explain the rise in default rates. It would be interesting to extend my framework in future research such that one can analyze how contract characteristics affect default rates within cohorts.

## 8 Analysis of two Bailout Policies

This section discusses an application of the presented structural model for policy analysis. I study a situation where the government is concerned about a destabilization of the financial system due to the losses that mortgage lenders incur from mortgage default. Assume that the government decides to neutralize all these losses by a suitable bailout policy. The question is then: Should the government bail out lenders or homeowners?

In case lenders are bailed out the government needs to cover the negative equity of defaulters, i.e. by how much the outstanding mortgage balance exceeds the value of the collateral. In contrast, the government could also give subsidies to homeowners who would otherwise default such that they continue to service the mortgage. This policy might well be cheaper because homeowners are willing to accept some negative equity and thus bear some of the losses on the house value unless they face severe liquidity problems. The subsidies then only have to overcome the temporary liquidity shortage to neutralize the losses for lenders. However it is also possible that subsidizing homeowners simply delays default to a later period such that the subsidy policy ends up being more expensive in the long-run. These opposing effects make a quantitative analysis desirable.

The two policies are compared by calculating the average cost per borrower who would default in absence of an intervention. For the bailout of lenders this simply amounts to the average negative equity of a defaulter which can readily be computed during the simulation. For the subsidy to homeowners one needs to modify the standard simulation procedure. Each period default decisions of borrowers given their liquid wealth, negative equity and other state variables are determined. Then for each potential defaulter the

subsidy required to make the borrower stay in the house is computed. When doing this the standard policy functions are used. This means borrowers will consume out of the subsidy, but further negative incentive effects are ruled out. The total sum of all subsidies to a cohort is divided by the number of defaulters without any intervention to make it comparable to the other bailout policy. The required real payment streams of both policies are compared by calculating present discounted values using the real interest rate  $r$ .

In order to account for the delayed default effect of the subsidy policy it is important to follow a cohort up to the point where the model does not predict any more default. Therefore this analysis will only be done for the 2002 cohort with the longest time horizon. Of course, this calculation can only be as accurate as the model captures actual default behavior. Since by construction the model explains the 2002 cohort relatively well this is an additional reason to focus on it. I find that bailing out lenders implies average real present discounted costs of 5.82% of the initial house price per borrower who defaults. In contrast subsidizing homeowners on average only costs 0.52% of the initial house price in real present discounted value terms. Bailing out lenders is thus 11 times more expensive than subsidizing homeowners. This is a huge difference.

A couple of comments on these results are in order. First, these are partial-equilibrium results. But it seems that general equilibrium effects of subsidizing homeowners would also be more favorable because keeping borrowers in their houses avoids downward pressure on house prices due to foreclosure sales. Second, both homeowners and lenders would probably prefer the subsidy to homeowners because borrowers like to stay in their houses and lenders do not have to deal with foreclosures and housing sales which will cause additional administrative costs for them. Finally, in reality one would of course need to take negative incentive effects into account. While both policies have negative incentive effects on lenders, the bailout of homeowners would also have negative incentive effects on borrowers. There might also exist practical problems of implementing a subsidy to homeowners in a fashion as assumed here. But one feasible policy could be to increase unemployment benefits for unemployed mortgage borrowers during a mortgage crisis such that they have enough resources to continue their mortgage payments. In any case these calculations show that there is potential for improving on policies that simply bail out the lenders both in terms of costs to taxpayers, but possibly also in terms of what lenders and borrowers would prefer.

## 9 Extension to lower Loan-to-Value Ratios

So far the paper focussed on loans with a LTV above 95% because these borrowers should be least likely to have a second mortgage on their home, cf. the discussion in section 2.1. The question arises whether the results of the paper also generalize to loans with a lower LTV. This section provides some evidence on this by repeating the reduced-form analysis of section 3 for loans with a LTV of the first mortgage between 75% and 84%. Due to the discussed data problems this section is necessarily somewhat tentative. Nevertheless, some very interesting results emerge.

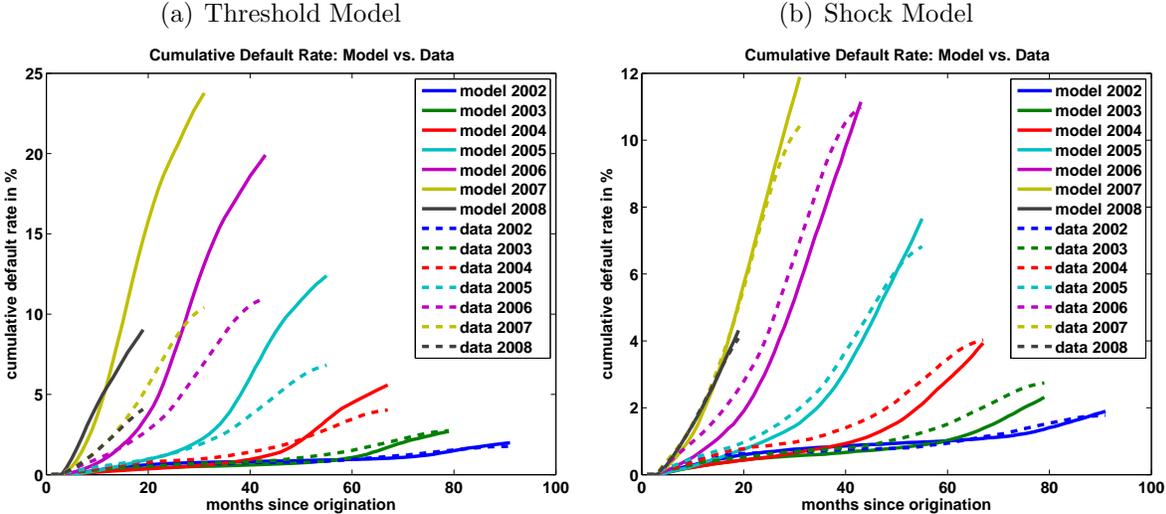
First I take the data for the loans with a LTV of the first mortgage between 75% and 84% at face value and assume that no one has a second mortgage. Accordingly the LTV varies within cohorts in steps of one percentage point between 75% and 84%. Changes to the distribution of loans over this support across cohorts observed in the mortgage data is again taken into account. The mortgage rate is again kept constant within a cohort and set equal to the respective cohort average. When estimating the models on the 2002 cohort I find that neither of the two models can capture this data well. Both models undershoot the cumulative default rate even for the most extreme parameter values where  $\phi = 0$  and  $\psi = 1$ . The reason is that the equity buffer generated by the down-payment is substantial for these borrowers. Because the 2002 cohort faced strongly increasing average house prices immediately after origination, too few borrowers in the simulation experience negative equity compared to observed default rates. It is important that both models fail if we take this data at face value. One can draw two possible conclusions from these results. Either we need a completely new theory of default for these loans or it is crucial to take second mortgages into account. I present evidence on the second explanation next.

Elul, Souleles, Chomsisengphet, Glennon, and Hunt (2010) report that 26% of all borrowers have a second mortgage and this adds on average 15% to the combined LTV. But they neither report a break-down of these statistics by the LTV of the first mortgage nor when borrowers take out the second mortgage. Faced with this situation I model a very simple form of intra-cohort heterogeneity taking these estimates of the frequency and size of second mortgages into account. I assume that 74% of borrowers have only one mortgage with a distribution of LTVs as in the mortgage data. But 26% of borrowers in each cohort independently of the LTV of the first mortgage also have a second mortgage adding 15% to the combined LTV. This implies that the support of the LTV distribution is expanded and also includes values between 90% and 99%. It is assumed that borrowers

got the second mortgage at the same time as the first one and pay the same mortgage rate on both. Admittedly, these are very crude assumptions. This exercise can only provide preliminary evidence until better data is available and should be regarded with considerable caution.

For this setup the reduced-form models are estimated again on the 2002 cohort. This yields estimates of  $\phi = -7.7\%$  and  $\psi = 2.4\%$ . The estimated models are again tested on their ability to predict out-of-sample. Figure 9 presents the results for all cohorts. The threshold model overshoots the data again. In contrast, the shock model provides an excellent fit to the data. Thus the double-trigger theory also provides a better explanation for this data under the maintained assumptions on second mortgages. Due to the discussed data problems I would personally put a lower weight on these results compared to the benchmark results. But these results suggest that the conclusions on the relative merit of the two theories are similar for this data.

Figure 9: Reduced-form results for borrowers with a first mortgage LTV of 75 – 84% taking second mortgages into account



## 10 Conclusions

This paper has presented simulations of theoretical default models for the observed path of aggregate house prices and a realistic microeconomic distribution. Theoretical predictions were then compared to data on default rates on prime fixed-rate mortgages to assess the explanatory power of the theories.

A test has been developed that examined whether estimated reduced forms of the frictionless option model and the double trigger hypothesis are able to predict out-of-sample. This test revealed that the frictionless default theory is too sensitive to the mean shifts in the house price distribution observed in recent years. In contrast, the double-trigger hypothesis attributing default to the joint occurrence of negative equity and a life event is consistent with the data.

Based on this finding a structural dynamic stochastic model with liquidity constraints and unemployment shocks was presented to provide micro-foundations for the double-trigger hypothesis. In this model the liquidity problems associated with unemployment can act as a trigger event for default. Accordingly, the level of negative equity at which individual borrowers default on their mortgage depends on non-housing state variables: liquid wealth and employment status. The model is broadly consistent with the data and explains most of the rise in mortgage default rates as a consequence of aggregate house price dynamics.

The structural model was used to analyze two bailout policies in a mortgage crisis. This revealed that in order to neutralize losses for lenders subsidizing homeowners is much cheaper than bailing out lenders when liquidity problems are a key determinant of mortgage default. A related policy question to which the model can be applied is how the design of unemployment insurance can help to prevent mortgage default.

The results of the reduced form and structural model as well as further supporting evidence on loan characteristics show that mortgage default has a strong macroeconomic component resulting from aggregate house price dynamics. This suggests caution to attribute the recent events entirely to a deterioration of loan quality. Instead, they hint at the existence of systematic macroeconomic risk in the mortgage market.

An important goal for future research is to develop an explanation of the house price boom and bust and the mortgage crisis in general equilibrium. This paper has presented a model where default rates match the data reasonably well taking house prices as given. It remains to provide a model that matches house prices as well as quantities in the housing and mortgage market. Obviously this represents a great challenge. But the model presented here may serve as a building block for that more general model.

## References

- Amromin, G. and A. L. Paulson (2009). Comparing Patterns of Default among Prime and Subprime Mortgages. *Federal Reserve Bank of Chicago Economic Perspectives Q2*, 18–37.
- Bajari, P., S. Chu, and M. Park (2010). An Empirical Model of Subprime Mortgage Default from 2000 to 2007. *Working Paper*.
- Calhoun, C. A. (1996). *OFHEO House Price Indexes: HPI Technical Description*. Washington, D.C.: Office of Federal Housing Enterprise Oversight.
- Campbell, J. Y. and J. F. Cocco (2003). Household Risk Management and Optimal Mortgage Choice. *Quarterly Journal of Economics* 118, 1449–1494.
- Campbell, J. Y. and J. F. Cocco (2011). A Model of Mortgage Default. *Working Paper*.
- Carroll, C. D. (1997). Buffer-Stock Saving and the Life-Cycle/Permanent Income Hypothesis. *Quarterly Journal of Economics* 112(1), 1–55.
- Corradin, S. (2009). Household Leverage. *Working Paper*.
- Davis, M. A., A. Lehnert, and R. F. Martin (2008). The Rent-Price Ratio for the Aggregate Stock of Owner-Occupied Housing. *Review of Income and Wealth* 54(2), 279–284.
- Deaton, A. (1991). Saving and Liquidity Constraints. *Econometrica* 59(5), 1221–1248.
- Demyanyk, Y. and O. Van Hemert (2011). Understanding the Subprime Mortgage Crisis. *Review of Financial Studies* 24(6), 1848–1880.
- Deng, Y., J. M. Quigley, and R. Van Order (2000). Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica* 68(2), 275–307.
- Elul, R., N. S. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010). What ‘Triggers’ Mortgage Default? *American Economic Review Papers and Proceedings* 100(2), 490–494.
- Foote, C. F., K. S. Gerardi, L. Goette, and P. S. Willen (2008). Subprime Facts: What (We Think) We Know about the Subprime Crisis and What We Don’t. *Federal Reserve Bank of Boston Public Policy Discussion Paper No. 08-2*.
- Foote, C. F., K. S. Gerardi, L. Goette, and P. S. Willen (2009). Reducing Foreclosures: No Easy Answers. *NBER Macroeconomics Annual* 24, 89–138.
- Foote, C. F., K. S. Gerardi, and P. S. Willen (2008). Negative Equity and Foreclosure: Theory and Evidence. *Journal of Urban Economics* 64(2), 234–245.
- Frederick, S., G. Loewenstein, and T. O’Donoghue (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* 40(2), 351–401.
- Gerardi, K. S., A. Lehnert, S. M. Sherlund, and P. S. Willen (2008). Making Sense of

- the Subprime Crisis. *Brookings Papers on Economic Activity* (Fall), 69–159.
- Gerardi, K. S., A. H. Shapiro, and P. S. Willen (2007). Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures. *Federal Reserve Bank of Boston Working Paper No. 07-15*.
- Ghent, A. C. and M. Kudlyak (2010). Recourse and Residential Mortgage Default: Theory and Evidence from U.S. States. *Federal Reserve Bank of Richmond Working Paper No. 09-10R*.
- Guiso, L., P. Sapienza, and L. Zingales (2011). The Determinants of Attitudes towards Strategic Default on Mortgages. *Working Paper*.
- Guvenen, F. (2006). Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective. *Journal of Monetary Economics* 53(7), 1451–1472.
- Jagtiani, J. and W. W. Lang (2010). Strategic Default on First and Second Lien Mortgages During the Financial Crisis. *Working Paper*.
- Kau, J. B., D. C. Keenan, and T. Kim (1994). Default Probabilities for Mortgages. *Journal of Urban Economics* 35, 278–296.
- Magnac, T. and D. Thesmar (2002). Identifying Dynamic Decision Processes. *Econometrica* 70(2), 801–816.
- Mayer, C., K. Pence, and S. M. Sherlund (2009). The Rise in Mortgage Defaults. *Journal of Economic Perspectives* 23(1), 27–50.
- Mian, A. and A. Sufi (2009). The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *Quarterly Journal of Economics* 124(4), 1449–1496.
- Quercia, R. G. and M. A. Stegman (1992). Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research* 3(2), 341–379.
- Vandell, K. D. (1995). How Ruthless is Mortgage Default: A Review and Synthesis of the Evidence. *Journal of Housing Research* 6(2), 245–264.
- Zeldes, S. (1989). Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence. *Quarterly Journal of Economics* 104(2), 275–298.