Mortgage calculator

Models - MA7404

course project

Sila Jan - js822@student.le.ac.co.uk

University of Leicester

December 21, 2015

Abstract

The project aims to develop a full functioning retail mortgage web-application calculator. The calculations are based on the underlying Markov Chain model that models and predicts the interest rate benchmark set quarterly by the Bank of England. In order to design the application, R software with Shiny package was used.
The app is hosted at http://jansila.shinyapps.io/Models_final and the code, including the dataset, is accessible at:
https://github.com/jansila/Models_mortgage.

Contents

1 Introduction

Interest rate is one of the fundamental variables throughout the history of economies. At the moment, the decision on the base rate in an economy, is in most countries left upon a central bank committee. They deliver their decision every three months.
That would be linked closely to their idea of the economy's dynamics. Then interest rate serves as one of the tools of monetary policy to steer the economy to reach primary goals set by the central bank - usually an inflation target. However, the rates have profound implications to the whole economy, not just the mortgage market, and directly influence, for instance, also exchange rates.

In the literature, the authors are not that concerned about the decision of the central bank. They usually try to model the short-term interest rates, that change within periods of days, or hours, not a benchmark that is constant for a quarter, which is a very discreet variable.

The continuous time is more applicable, and key, in pricing securities. An important conclusions in this area are presented by Gray (1996), where a Generalized Autoregressive Conditional Heteroskedasticity model (GARCH) is used to model the rate with regime switching governed by a Markov process. As it is trying to capture the changes in continuous time, Brownian motion process lies in the heart of the model, just as it does in Vasicek (1977). Volatility of the short term interest rates is captured with the time-varying volatility family of model (ARCH), introduced by Engle (1982).

In our analysis, that will lead to setting up an web application for mortgage quotes, we shall restrict ourselves to forecast the discreet rates that are announced by the Bank of England. For that, continuous time models are not suitable, or rather necessary. Hence, we shall model the interest rate changes with Markov Chain.

The literature in this particular field does not seem to be existing at all. As mentioned earlier, most of it is concerned with pricing derivatives and bonds, or uses Markov Chains in a rather complex setups. Some references can be made to Kijima & Komoribayashi (1998), simulations with Markov Chain are used in Norberg (1995). However, we will have to rely on precise definition of the process for our purposes, as the literature does not provide much guidance in terms of application.

2 Methodology

The underlying method for modeling the interest rate is *Markov chain* in discrete time with discrete state space. The method is used in this project as introduced in Cox & Miller (1977).

Cox defines Markov chain as a sequence $X_0, X_1, ...$ of discrete random variables with the property that the conditional distribution of X_{n+1} given $X_0, X_1, ..., X_n$ depends only on the value of X_n, but not further on $X_0, ..., X_{n-1}$.

Definition 1. *For any set of values h,j, ..., k belonging to the discrete state space,*

$$prob(X_{n+1} = k|X_0 = h, ..., X_n = j) = prob(X_{n+1} = k|X_n = j).$$

The above mentioned definition is called *Markov property* and is the corner stone of the method. It is easily deducible, that any process with independent increments possesses the Markov property, Grechuk (2012).

2.1 Markov chains for interest rates

The Markov chain family of models is suitable for this type of problem as it relevant not only mathematically, but also economically. In most of the developed world, every three months, a board of governors meets to discuss and choose the interest rate benchmark for their country, be it Federal Reserve System (Fed) in the United States, European Central Bank (ECB), or Bank of England (BoE). As they can set the rates to an arbitrary value, the decision every quarter is independent of the last time. In other words, if the central bankers wished 5% rates, they don't need to raise then twice by 2.5% each time.

Hence, keeping in mind the result mentioned from Grechuk (2012), it would make sense to model the interest rate with Markov chain, that has three assumptions, Grechuk (2012);

Markov chain assumptions:

- The state space is *discrete* – only 44 unique value over 105 years of history, one every three months
- The values (and also their increments) can be considered to be independent and are clearly *countable*
- The Markov property holds
- We further assume the chain to be time-homogenous in its nature

Looking at the way the interest rates are determined, we see the relevance of this particular method to describe, and forecast, them. However, we should formally check the Markov property, as the other two aforementioned assumptions are clearly satisfied.

It is also possible, and for the sake of simplicity of the model, to use differenced data.
Firstly, the original data contain an absorbing barrier. The current state of historically lowest interest rates of 0.05% is unprecedented in history. Therefore, the process would always remain in the state, once it would be achieved. And since it is the current state, the model would never leave it. We, however, do not in this case lose any statistical or economical relevance.
Secondly, the state space shrinks from 44 to 23 states, which makes it more comprehensible.
Lastly, the literature, although concerned more with continuous time diffusion, suggests this approach.

2.1.1 The Markov property

Despite the innumerable studies in financial economics rooted in the Markov property, there are only two tests available in the literature to check such an assumption, according to Chen & Hong (2012). It seems that the current literature is mostly concerned with non-parametric procedures as described in Aït-Sahalia (2000) and Aït-Sahalia (2002). To build a nonparametric testing procedure, the first test uses the fact that the Chapman-Kolmogorov equation must hold in order for a process to be compatible with the Markov assumption, in detail in Aït-Sahalia (2002).
Those are, however, used for more complicated models on stock markets with much richer dataset structure. For our purposes, it is sufficient to construct the test as described in Grechuk (2012).

Definition 2. *For a set of observations $x_1, x_2, ..., x_N$, let n_{ijk} be the number of times t, such $(1 \leq t \leq N-2)$, such that $x_t = i, x_{t+1} = j$ and $x_{t+2} = k$. Then the test statistic:*

$$X^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - n_{ij}\hat{p}_{jk})^2}{n_{ij}\hat{p}_{jk}}$$

is approaching the χ^2 distribution with $r = |\mathrm{S}|^3$ degrees of freedom.

2.1.2 Transition probabilities

Since we assume our model to be time homogenous, that implies a stationary transition mechanism that will be described later, we can introduce *transition probabilities* as defined in Cox & Miller (1977) and Grechuk (2012).

Definition 3. *For a homogenous Markov chain, we define the n-step transition probabilities as*

$$p_{jk}^{(n)} = prob(X_{m+n} = k|X_m = j) = prob(X_n = k|X_0 = j), (m, n = 1, 2, ...).$$

Where individual 1-step probabilities are estimated as:

$$\hat{p}_{ij} = \tfrac{n_{ij}}{n_{i\cdot}}$$

, where n_{ij} are transitions from state i to j and $n_{i\cdot}$ are all transitions from i.

2.1.3 Irreducibility and periodicity

An irreducible chain is one in which all states inter-communicate, Cox & Miller (1977). That means, that it is possible, in a finite number of steps, reach any state regardless of the origin. That implies that the chain forms a single closed set and all its states are of the same type, i.e. none of the states are absorbing.

Also, it may be the case, that we always come back to a certain state after some number of steps. In that case, we speak about periodic chains. Formal definition as in Spedicato (2015).

Definition 4. *A state s_i has period k_i if any return to state s_i must occur in multiples of k_i steps. That is*

$$k_i = gcd\{n : Pr(X_n = s_i | X_0 = s_i) > 0\},$$

where gcd is the greatest common divisor. If $k_i = 1$ the state s_i is said to be aperiodic, else if $k_i > 1$ the state s_i is periodic with period k_i.

If, on top of that, if the chain is aperiodic, it is called *ergodic*, Cox & Miller (1977). That implies the existence of unique row vector π called the *equilibrium distribution*.

Definition 5. *The equilibrium distribution is defined as follows:*

$$\lim_{n\to\infty} P^{(n)} = \lim_{n\to\infty} P^{(0)} P^{(n)} = P^0 \begin{pmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{pmatrix} = \pi$$

which we will call *stationary distribution* in our case.

3 Data and results

The data is the *British interest rate*, also known as Bank of England *base*, as set by the Bank of England since 1911. The dataset comprises of quarterly figures, as announced by the bank starting with 1911 Q1. Hence there are 417 observations and the data were downloaded from the Bloomberg terminal.

As far as the time period is concerned, the whole sample was included as even almost hundred years old data are relevant for today. For example, World War II period represents a deep economical recession, just as we have witnessed in the past decade. That enables us to inspect the behavior of the interest rate in such a time and makes the current state non-extreme. Moreover, we can see the recovery, that, our economy seems to be on the track today.
The whole sample also provides reference to all states of the economy, from aforementioned recession, to the booming 1980's and 1990's where the UK experienced double digits rates for almost two decades.

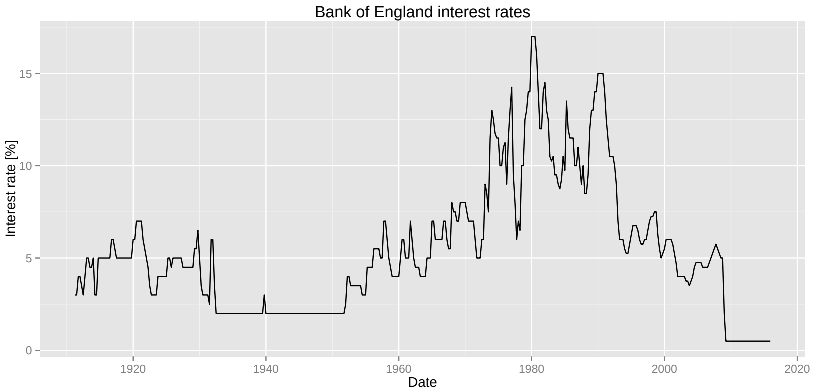

Figure 1: Time series of the Bank of England interest rate benchmark
Source: Data: Bloomberg, graph: author's computation

However, as explained in more detail in section 2, the data were differenced. That yielded 416 observations that represent interest rate hikes and cuts.

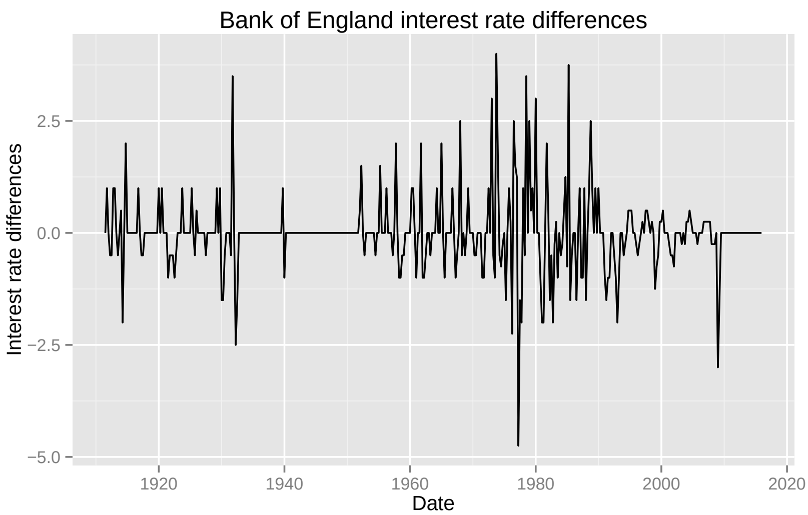

Figure 2: Bank of England interest rate benchmark hikes and cuts
Source: Data: Bloomberg, graph: author's computation

We observe behavior similar to market volatility – interchanging periods of high and low volatility and its clustering. As discussed above, last 5 years the rates are at constant 0.05%, hence the graph shows zero as in the WWII period.

Firstly, the data need to evaluated according to the assumptions and properties stated in section 2.

3.1 Model appropriateness and estimation

3.1.1 Markov property

Although the package *markovchain* used in the model, created by Spedicato (2015), has proved itself to work very well, it does not offer the test statistic needed. Since the precise test statistic to check the Markov property, that is presented above, was not available in any R package or anywhere else, I constructed it myself.

The software of my choice for this whole project is R, by R Core Team (2015). Particularly for the possibility to create an actual web application with Shiny by Chang *et al.* (2015) later on.

The function was thoroughly tested against series that should possess the property and those that should not.

The code can be found in the Section 8 and since it is not a necessary part of the web application, it has been uploaded to a GitHub repository **checkMP** to develop further.

The method called *checkMP* returns output that shows the value of the test statistic, and results of the test with the null hypothesis H_0 :*the series has Markov property.*

It shows if the Markov property is to be rejected for the 95% confidence level – $rejection[1,1] and the 99% confidence level – $rejection[1,2]

	checkMP output	
	$test_statistic	
[1]	885.1531	
	$rejection	
	[, 1]	[,2]
[1,]	FALSE	FALSE

Table 1: Output for checking the Markov property on our sample

The original, undifferenced data were tested as well and still, the Markov property was not rejected either. The value of the test statistic in that case is 529.8 and at neither level the H_0 is rejected.

3.1.2 Irreducibility and periodicity

Let us inspect the property of discreteness of the state space, the irreducibility and periodicity of the model in question.

According to the package *markovchain* for R, the time series is irreducible and all the states share common periodicity 1, as presented in the model output in Figure 4. That means the model is aperiodic – no state is systemically achieved after certain amount of steps again.

We can also inspect it, graphically. Of course, at this point, we are assuming it is at all possible. Also to enhance readability, the transition probabilities were hidden.

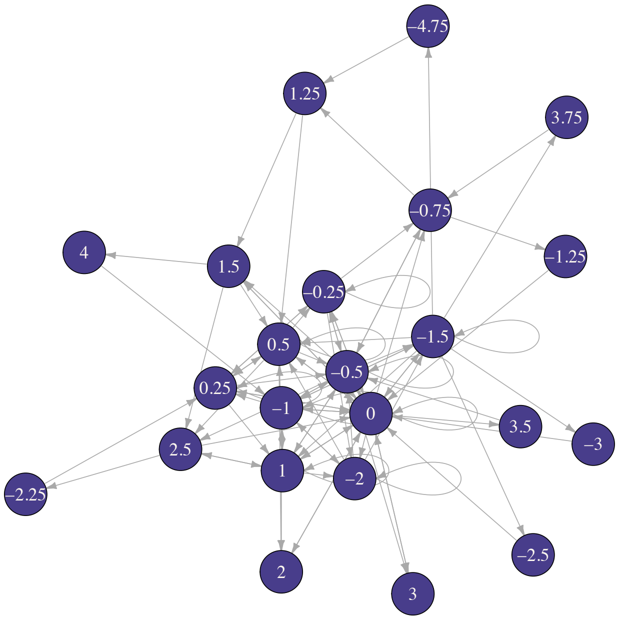

Figure 3: Graphical representation of the chain
Source: Output by Spedicato (2015) and Csardi & Nepusz (2006)

We see quite clearly, that the model has indeed the desirable properties; the state space is *finite* with 23 states; is *irreducible* with no absorbing, and all communicating, states and seems to be *aperiodic*. The last property is harder to show in the graph properly, but we can test it within the R package *markovchain*, that shows the periodicity to be 1.

3.1.3 Transition probabilities

Seeing the interest rate series satisfies the necessary conditions, we know the Markov chain model makes sense and therefore, we estimate it.

The summary of the model is as follows:

```
MLE Fit  Markov chain that is composed by:
Closed classes:
-0.25 -0.5 -0.75 -1 -1.25 -1.5 -2 -2.25 -2.5 -3 -4.75 0 0.25 0.5 1 1.25 1.5 2 2.5 3 3.5 3.75 4
Recurrent classes:
{-0.25,-0.5,-0.75,-1,-1.25,-1.5,-2,-2.25,-2.5,-3,-4.75,0,0.25,0.5,1,1.25,1.5,2,2.5,3,3.5,3.75,4}
Transient classes:
NONE
The Markov chain is irreducible
The absorbing states are: NONE
```

Figure 4: Summary of the Markov chain model
Source: Output by Spedicato (2015)

Let us present the transition probabilities' estimates.[1] The initial state of the model is highlighted in gray to see instantly the next possible steps and for better orientation between negative and positive changes:

	-0.25	-0.5	-0.75	-1	-1.25	-1.5	-2	-2.25	-2.5	-3	-4.75	0	0.25	0.5	1	1.25	1.5	2	2.5	3	3.5	3.75	4
-0.25	0.154	0.231	0.077	0	0	0	0.077	0	0	0	0	0.385	0.077	0	0	0	0	0	0	0	0	0	0
-0.5	0.054	0.189	0.027	0.135	0	0.081	0	0	0	0	0	0.405	0	0	0.027	0	0.027	0	0.027	0.027	0	0	0
-0.75	0	0.500	0	0	0.250	0	0	0	0	0	0	0	0	0	0	0.250	0	0	0	0	0	0	0
-1	0	0.136	0	0.227	0	0.045	0.045	0	0	0	0	0.364	0.045	0	0.091	0	0	0.045	0	0	0	0	0
-1.25	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
-1.5	0	0	0	0.091	0	0.091	0	0	0.091	0.091	0.091	0.182	0	0.091	0.182	0	0	0	0	0	0	0.091	0
-2	0	0.167	0	0.333	0	0.167	0.167	0	0	0	0	0	0	0.167	0	0	0	0	0	0	0	0	0
-2.25	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
-2.5	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
-3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
-4.75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	0.025	0.066	0.004	0.029	0	0.021	0.008	0	0	0	0	0.710	0.017	0.012	0.075	0	0.008	0.012	0	0.004	0.008	0	0
0.25	0.133	0.067	0	0	0	0	0	0	0	0	0	0.200	0.400	0.133	0.067	0	0	0	0	0	0	0	0
0.5	0.077	0.077	0	0	0	0	0	0	0	0	0	0.308	0.154	0.231	0	0	0	0.077	0.077	0	0	0	0
1	0	0.037	0	0.037	0	0	0.037	0	0	0	0	0.741	0	0.037	0.074	0	0	0	0.037	0	0	0	0
1.25	0	0	0	0	0	0	0	0	0	0	0	0	0	0.500	0	0	0.500	0	0	0	0	0	0
1.5	0	0	0	0	0	0	0	0	0	0	0	0.250	0	0.250	0	0	0	0	0.250	0	0	0	0.250
2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
2.5	0	0	0	0	0	0	0	0.250	0	0	0	0.500	0	0	0.250	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
3.5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3.75	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5: Transition probabilities' estimates
Source: Computed by Spedicato (2015)

The table shows that being in the state 0, that has occurred exactly 242 times out of 416 observations, the system is most likely to stay in it with probability 0.71.

[1]Please note, that the rows might not necessarily sum up to 1, as the figures were rounded for the sake of readability.

3.2 Forecasting the interest rate

Firstly, it is important to see, if the forecasts are relevant and exhibit similar behavior as the underlying historical data. In our case, they do, especially when forecasting the interest rate changes. It often produces constant segments just as the historical data and jumps up and down. Although, with no expert in the field available, we cannot perform a ”proper” Turing test. However, we can confidently conclude, that the forecasts look very much alike, at least in a sizable majority of the time. We also compare more simulations in Figure 7.

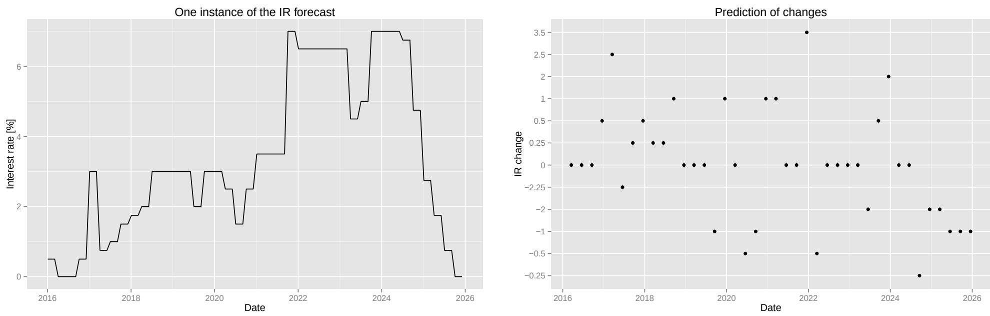

Figure 6: Example of forecasts
Source: Author’s computations

We are aware, that the system will, after a large number of iterations, converge to stationary distributions for each state. Even though, for our purposes, we shall continue with our forecast in the quarterly innovations, the number of iterations in the following few years is not as high as to motivate us to use the stationary distributions. Also, solely from the computational point of view, the calculations are not that demanding to force us to use the stationary distributions instead of following the probability paths of the Markov chain.

Let us, however, include them in a table below:

Table 2: Stationary distributions

state	-0.25	-0.5	-0.75	-1	-1.25	-1.5	-2	-2.25	-2.5	-3	-4.75
prob	0.031	0.089	0.010	0.053	0.002	0.027	0.014	0.002	0.002	0.002	0.002
0	0.25	0.5	1	1.25	1.5	2	2.5	3	3.5	3.75	4
0.581	0.036	0.031	0.065	0.005	0.010	0.012	0.010	0.005	0.005	0.002	0.002

Source: Author's computations, Spedicato (2015)

The distribution also shows, that is slightly skewed towards negative values. In the long run, a negative change (interest rate cut) would happen in $23,6\%$ cases, no-change in 58% cases and interest hike only in 18% cases. Therefore, its is quite likely, the model might end up very deeply in the negative territory, which, in reality cannot really happen. Although, they are in practice possible as shows the latest example on ECB's deposit facility, we disregard them.[2]

The purpose of negative rates is beyond this text, but since it is mostly used in inter-bank loans, or the bond market, we disregard this possibility. We can assume, that a mortgage, as a retail bank product would never be issued under such circumstances. Hence, the underlying interest rate is not allowed to drop below -0.5%. To illustrate, I simulated 100 time series of the base interest rate forecast.

[2] https://www.ecb.europa.eu/press/pr/date/2015/html/pr151203.en.html

Let us take a closer look at how the forecasts behave:

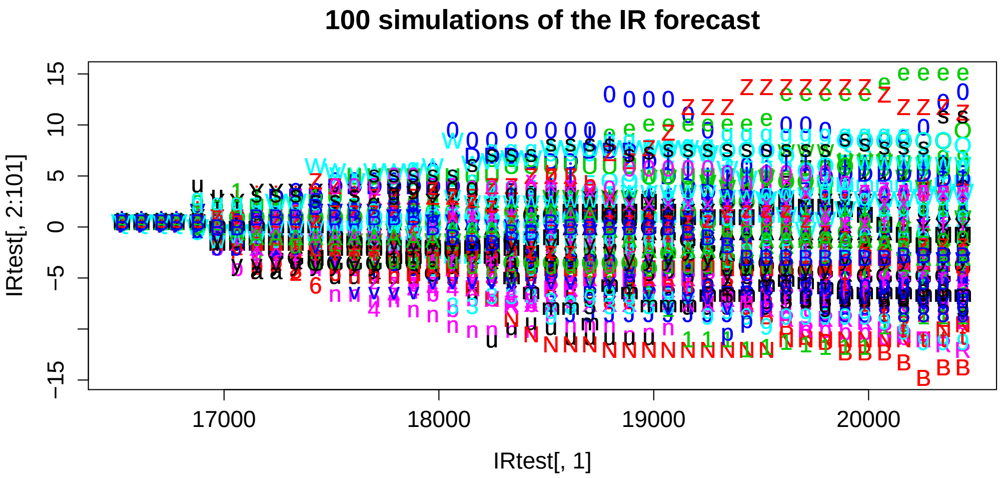

Figure 7: Plot of 100 simulations of 10yr interest rates
Source: Author's computations

We see the variance of the forecast growing substantially. Since the bank's markup above the threshold can hardly be around 15 percentage points to break-even, we adjust the rate to never be less than negative 0.05. Currently the mortgage rates are between 3 to 4 percents. Hence, I turned to the above mentioned adjustment for the sake of the model's relevance.

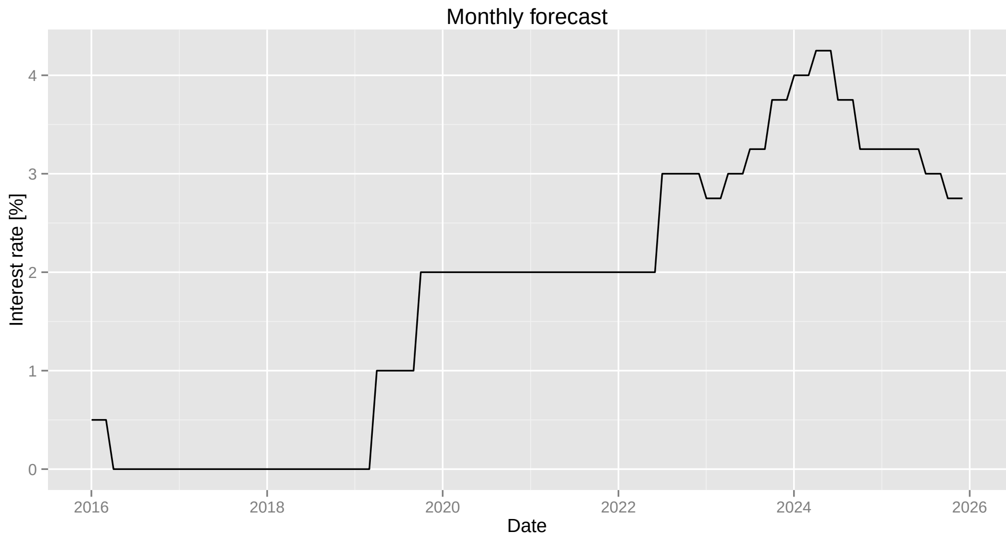

Figure 8: Monthly adjusted interest rates prediction - example
Source: Author's computations

Lastly, in order to be able to compute the interest on a monthly basis, the forecast is transformed, so that each new rate holds for a total of three months, until can be changed again. One example of the final series is depicted above. Also, let us interpret this rate as year rate. In other words, the rate for each month means how much would one pay for a year long contract from that month.

4 Mortgage calculator

The calculation of the mortgage payments, of course relies on key parameters such as the length of the contract, the upfront payment for the house and the underlying interest rate benchmark.

Firstly, we compute *Loan to Value* ratio defined as: $LtV = \frac{loan - deposit}{loan}$. That says how much is the customer willing/or able to deposit. It is directly linked

with their perceived riskiness - customers with no deposits would be more likely to default on their payments. Therefore, the bank's interest rate margin will penalise this ratio.
Also, if the customer is buying their first property, it is expected to have higher probability of default, therefore will be charged little more on the interest payments.
Finally, if the clients is already a customer, then we expect to know about them little more that would help us to evaluate their creditworthiness better. The mortgage interest rate is, therefore, calculated as follows:

Definition 6. *Let us define the mortgage interest rate as follows:*

$$\begin{aligned} interest\ rate_{mortgage} =& interest\ rate_{BoE} + 1.5 \cdot LtV + 2_{ifLtV>0.9} \\ & + 0.5_{1^{st}timer} - 0.1_{ifcustomer} \end{aligned} \tag{1}$$

Then the mortgage repayment is calculated using the traditional formula for fixed-rate mortgage amortization, but it is reevaluated every month, as the model works with variable interest rate. The formula is used as defined in Brown & Matysiak (2000).

Definition 7. *The fixed rate amortization (Month's purchase in n months) is defined as:*

$$MV = balance \cdot \frac{r(1+r)^n}{(1+r)^n - 1}$$

Where balance, is the remaining balance, r is the annualised monthly interest rate and $n \in \{0, \ldots, number\ of\ months\}$ is the number of remaining months.

Having computed the periodic payment, we also compute the month's *interest payment* simply as $r \cdot (remaining\ balance)$, then the *principal paid* is the difference of the month payment and respective interest payment. We subtract this value from the current balance. This process will eventually, amortize the loan, which means that the principal as well as the interest are paid off.

Repayment development of one instance is pictured here:

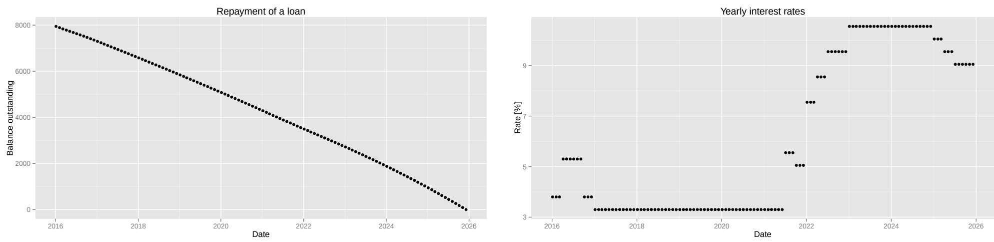

Figure 9: Mortgage repayment - example
Source: Author's computations

The above figure depicts repaying a mortgage on 10,000 with 2,000 deposit for 10 years for a first-time buyer. We see the interest rate to go up to 10% in the 2020's, That increases the monthly payments and in the figure below, that depicts the payment's allocation, that more goes towards paying off the interest and less towards the principal for a few years.

Figure 10: Mortgage repayment - payment structure
Source: Author's computations

Since the core of the model is stochastic, every evaluation of the model produces a new result. The model does not involve any fixations, or limitations, except for not allowing negative interest rates, as discussed above.
Hence, a result for the same set-up will be 'always' different. In other words, should a user reload the page, or set a different set of values, they will not get back to the same result.
Although in reality, this would not be permitted for it would be too confusing for the customer and probably even against regulation. Also, since the interest rate is not capped from above, theoretically, extremely large values are possible. Although the simulations, run in limited environment, did not produce such a result, it cannot be ruled out.
A decision was made not to cap it, as this is more of a mathematical, than business, exercise, and therefore was left untouched to clearly see the behavior of the model.
Let us remind, that the negative rates were set to zero, so there are no cases of high negative rates that do not make sense economically.

4.1 Sensitivity

As was mentioned earlier, the model of the interest rate, and thus the mortgage as well, is rather sensitive. It is sensitive to the contract length, since the variance tends to grow in time. The interest rates vary in the interval (-60, 60) when considering 50 years and more.
However, it hardly got over the rate of 80 each way, even for much longer series; i.e. 500 years or more. Hence, the model seems relatively balanced. Although, it is necessary to point out, that those interest rates are hardly imaginable in there real world. Should the application be used in real life, it would be logical to cap the interest costs. A solution would be to choose

a set of random numbers (in R set.seed(x) function) that would model the rates in line with the company's expectations, even for the longest series. In this instance, it was decided not to adopt this approach to clearly show the model behavior and its flaws.

5 Web application

The ultimate goal of this project is to design a web application. I chose to use statistical software R, by R Core Team (2015), that would serve as the engine for an app developed with Shiny package by Chang *et al.* (2015). The app is accessible at hosting on shinyapps.io and all the code is published in its own repository on GitHub.

As far as using and testing the application is concerned, it is possible to run it locally on a computer with R distribution, that can be freely obtained at its home site: `https://www.r-project.org`. Much advised is to use an IDE by RStudio Team (2015)[3]s which was also used in developing this project.
Alternatively, as mentioned, it can be accessed on shinyapps.io, however, the hosting there allows only 25 active hours per application with author's current account. Therefore, the application might not render later.

5.1 Graphical User Interface

The application is inspired by the current market standard set by High Street banks such as Lloyds, Barclays, NatWest and the RBS. The intention was to include enough parameters to quote a price for the mortgage, but not to put

[3] `http://www.rstudio.com/`

too much pressure on the user.
Therefore, the GUI was kept simple, with a clear layout: inputs in the left hand side panel and outputs on the right hand side.

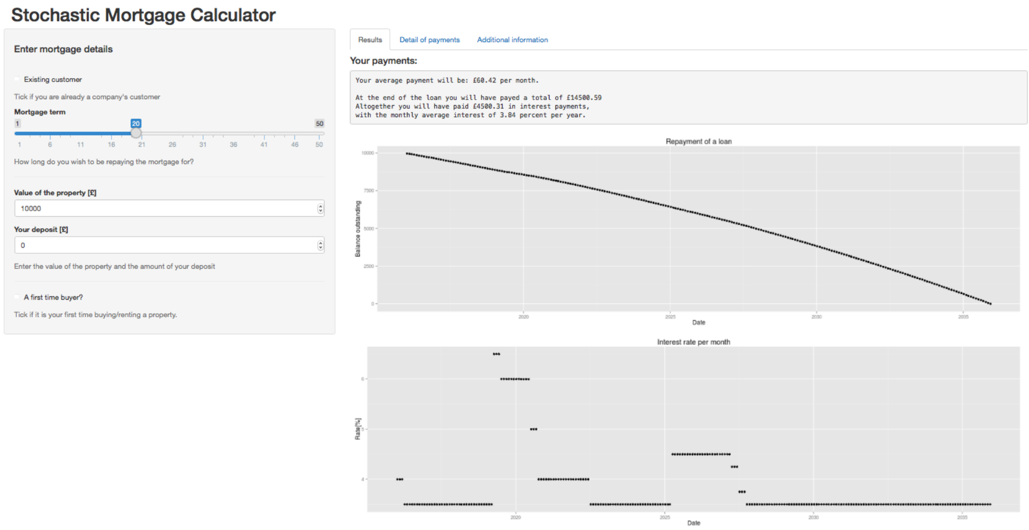

Figure 11: Graphical User Interface

The main part of the GUI has three tabs:

Results - default screen which shows the **Your payments** text box. That summarizes the mortgage amortization calculations, particularly average monthly payment. Also, it shows the total payment and how much **interest payment** will be. It also includes the average yearly interest rate. This tab also includes a graphical representation of the amortization of the loan. In other words, it shows outstanding loan at any month. It is represented by points, hence it might seem, that it drops below zero, but it does actually not.

Detail of payments - second tab plots the allocation of each months payment. Hence, it shows the development of the month's payments over time and how much of that goes towards the principal and paying off the interest. Example is in Figure 10.

Additional information - last section briefly describes the model and its functionality. It also states the references and provides links to the codes and project documentation.

5.2 User inputs

The user is asked to input few parameters. Please, refer to Figure 11:

- Existing customer - a boolean value if the customer is already company's client. True value gives a discount on the interest rate.
- Mortgage term - user inputs on a slider how long they want to repay the mortgage for. It takes values from 1 to 50 years by 1 year.
- Value of the property - states how much does the desired property costs.
- Your deposit - how much money the costumer pays upfront. Subtracted from the Value of the property, it gives the *principal* of the loan.
- First time buyer? - a boolean parameter that states if it is the customer's first property bought. If yes, it slightly raises the interest rate.

All of these parameters are directly used in calculation of the mortgage scheme. Its result is the quote on the loan costs for the user. The code that calculates is at the mentioned GitHub repository in file server.R.

6 Conclusion

The purpose of this project was to model the interest rate set by Bank of England and use it to design a web application to serve as a mortgage calculator. The method used to model rates was a discreet Markov chain. The model makes economical sense and fits well to the historical data. However, it might behave rather explosively, if consequent hikes occur, which is permitted. Whereas deep cuts into negative territory are manually prohibited and set to zero, if occur.

The ultimate goal was developing a web application that is hosted at shinyapps.io, which was achieved.
It could be further improved by addressing the extreme cases and possibly link with a real pricing policy of a bank or a building society.

7 Bibliography

References

Aït-Sahalia, Y. (2000): "Testing the markov hypothesis in discretely sampled continuous-time processes." *Technical report*, Working paper, Princeton University.

Aït-Sahalia, Y. (2002): "Telling from discrete data whether the underlying continuous–time model is a diffusion." *The Journal of Finance* **57(5)**: pp. 2075–2112.

Brown, G. R. & G. A. Matysiak (2000): *Real estate investment: A capital market approach*. Harlow: Financial Times Prentice Hall.

Chang, W., J. Cheng, J. Allaire, Y. Xie, & J. McPherson (2015): *shiny: Web Application Framework for R*. R package version 0.12.2.

Chen, B. & Y. Hong (2012): "Testing for the markov property in time series." *Econometric Theory* **28(01)**: pp. 130–178.

Cox, D. R. & H. D. Miller (1977): *The theory of stochastic processes*, volume 134. CRC Press.

Csardi, G. & T. Nepusz (2006): "The igraph software package for complex network research." *InterJournal* **Complex Systems**: p. 1695.

Engle, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica: Journal of the Econometric Society* pp. 987–1007.

Gray, S. F. (1996): "Modeling the conditional distribution of interest rates as a regime-switching process." *Journal of Financial Economics* **42(1)**: pp. 27–62.

Grechuk, B. (2012): "Module ma 7404 models." *Lecture notes* .

Kijima, M. & K. Komoribayashi (1998): "A markov chain model for valuing credit risk derivatives." *The Journal of Derivatives* **6(1)**: pp. 97–108.

Norberg, R. (1995): "A time-continuous markov chain interest model with applications to insurance." *Applied stochastic models and data analysis* **11(3)**: pp. 245–256.

R Core Team (2015): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

RStudio Team (2015): *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Spedicato, G. A. (2015): *markovchain: discrete time Markov chains made easy*. R package version 0.4.3.

Vasicek, O. (1977): "An equilibrium characterization of the term structure." *Journal of financial economics* **5(2)**: pp. 177–188.

8 Appendix

```r


checkMP<-function(series=ir_diff[,2]){
  require(markovchain)

  transMatrix<-markovchainFit(data=series)$estimate@transitionMatrix

  #make the n-2x3 matrix for observations
  subSample<-series[1:(length(series)-(length(series)%%3))]

  seqSet1<-matrix(c(subSample[1:(length(subSample)-2)],
                    subSample[2:(length(subSample)-1)],
                    subSample[3:(length(subSample))]
  ),ncol=3) #fill the matrix in reverse order so position 11 is the first obersvation,12 second and 13 third
  #compute row frequencies
  temp<-as.data.frame(seqSet1)
  Nijk<-aggregate(temp, by=temp, length)[1:(ncol(temp)+1)]

  seqSet2<-seqSet1[,-3] #make matrix of couples
  temp2<-as.data.frame(seqSet2)
  Nij<-aggregate(temp2, by=temp2, length)[1:(ncol(temp2)+1)] #rowfrequencies included

  findNijPjk<-function(Nijk=Nijk, Nij=Nij, trans=transMatrix, row=1){
    i<-Nijk[row,1]
    j<-Nijk[row,2]
    k<-Nijk[row,3]

    fromCh<-as.character(j)
    toCh<-as.character(k)
    Pjk<-trans[fromCh,toCh]

    m1<-which(Nij[,1]==i)
    m2<-which(Nij[,2]==j)
    m<-c(m1,m2)
    return(Nij[m[anyDuplicated(m)],3]*Pjk)
  }
  test<-c(length=dim(Nijk)[1])
  #compute the test statistic
  for(z in 1:dim(Nijk)[1]){
    foundNijPjk<-findNijPjk(Nijk=Nijk, Nij=Nij, trans=transMatrix, row=z)
    test[z]<-((Nijk[z,4]-foundNijPjk)^2)/foundNijPjk
  }
  result<-sum(test)
  #return value of the test statistic and test at confience level 95% and 99%
  return(list(test_statistic=result,rejection=cbind(result>qchisq(0.95,df=length(series)^3),result>qchisq(0.99,df=length(series)^3))))
}
```

checkMP.R